AD-A065 651 OHIO STATE UNIV RESEARCH FOUNDATION COLUMBUS F/G 1/3
USAF-ASEE (1978) SUMMER FACULTY RESEARCH PROGRAM (WPAFB). VOLUM--ETC(U)
NOV 78 C D BAILEY F44620-76-C-0052
UNCLASSIFIED AFOSR-TR-79-0232 NL

1 OF 5
AD
A065651

AFOSR-TR- 79-0232

LEVEL II

NW

AD A0 65651

AD 65 650

DDC FILE COPY

1978

USAF-ASEE SUMMER FACULTY RESEARCH PROGRAM WPAFB

D D C
MAR 14 1979
C

Volume II of II

Research Report

OSU

The Ohio State University
Research Foundation
Columbus, Ohio 43212

November 1978

79 03 12 013

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

18 REPORT DOCUMENTATION PAGE	READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER 19 AFOSR-TR-79-0232	2. GOVT ACCESSION NO. (1978)
	3. RECIPIENT'S CATALOG NUMBER
4. TITLE (and Subtitle) 6 USAF-ASEE Summer Faculty Research Program ~~(Volume II)~~ (WPAFB). Volume II.	5. TYPE OF REPORT & PERIOD COVERED 9 Final rept. 1 Jan 76 - 30 Sep 78,
	6. PERFORMING ORG. REPORT NUMBER
7. AUTHOR(s) Prof Cecil D. Bailey 10	8. CONTRACT OR GRANT NUMBER(s) 15 F44620-76-C-0052
9. PERFORMING ORGANIZATION NAME AND ADDRESS Ohio State University Research Foundation Columbus, OH 43212 17 D1	10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS 61102F 16 2307 D1
11. CONTROLLING OFFICE NAME AND ADDRESS Air Force Office of Scientific Research/XOP Bldg. 410 Bolling AFB, Washington D.C. 20332	12. REPORT DATE 11 November 1978
	13. NUMBER OF PAGES 215
14. MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office) 12 429p.	15. SECURITY CLASS. (of this report) UNCLASSIFIED
	15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

Approved for public release; distribution unlimited.

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

18. SUPPLEMENTARY NOTES

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

The fourth USAF-ASEE Summer Faculty Program was conducted at Wright-Patterson AFB during the summer of 1978. A total of twenty-five research associates (participants from sixteen universities in eleven states participated in the program. The program was again highly successful as judged from the participants' written reports of research accomplished, verbal contacts with those in the program (both participants and associated Air Force personnel) and upon written evaluations of the program by both the participants and Air Force personnel.

DD FORM 1 JAN 73 1473 267 360 UNCLASSIFIED LB

TABLE OF CONTENTS

Volume I

AF WRIGHT AERONAUTICAL LABORATORIES

AF AVIONICS LAB., Continued

Contents:

Volume II

AEROSPACE MEDICAL RESEARCH LABORATORY

AIR FORCE HUMAN RESOURCES LABORATORY

AIR FORCE LOGISTICS COMMAND

A

B

AEROSPACE MEDICAL RESEARCH LABORATORY

Research Associates:

Xavier J.R. Avula, University of Missouri -- Rolla

Lloyd F. Elfner, Florida State University

Duane G. Leet, University of Dayton

Stephen M. Strauch, The Ohio State University

Glenn F. Wilson, Wittenberg University

MATHEMATICAL MODELING OF THE RESPONSE OF THE VASCULAR SYSTEM TO TIME-DEPENDENT ACCELERATIONS

Xavier J.R. Avula*
University of Missouri-Rolla, Rolla, Missouri

ABSTRACT

In this study a mathematical treatment of the response of blood vessels to transient and linearly varying acceleration forces is presented. For the response of the large vessels, a numerical method based on solving the coupled partial differential equations derived from the Navier-Stokes equations and the theory of large elastic deformations is utilized under appropriate boundary conditions. The effect of acceleration on small blood vessels *per se* is shown to be negligible from an order of magnitude viewpoint. However, there is an indirect effect which is brought about by blood pooling in large vessels and the consequent changes in pressure gradient across the microvascular bed. In an effort to determine this pressure gradient, the arterial pressures are calculated for examples of linearly varying and transient $-G_z$ acceleration profiles, and one of the solutions is favorably compared with a pressure variation from an animal experiment reported in literature. The relationship between the pressure and flow rate in a narrow blood vessel is obtained under Stokes flow approximation. Since computation of venous pressure under acceleration forces is also necessary to include the microcirculation in the overall mathematical model, recommendations for further research are made.

*Professor, Department of Engineering Mechanics

MATHEMATICAL MODELING OF THE RESPONSE OF THE VASCULAR SYSTEM TO TIME-DEPENDENT ACCELERATIONS

Xavier J.R. Avula*
University of Missouri-Rolla, Rolla, Missouri

SUMMARY

A mathematical model to determine the response of blood vessels to time-dependent accelerations is presented. The model is based on the coupled partial differential equations derived from the Navier-Stokes equations and the theory of large elastic deformations. Using the model the pressures in the aorta is calculated for a linearly varying and a transient $-G_z$ acceleration. One of the solutions is compared favorably with an experimentally determined pressure in an animal experiment. The effect of acceleration on the flow in narrow blood vessels is shown to be negligible from an order of magnitude viewpoint. Since the blood pooling in large vessels indirectly affect the flow through the microvascular bed a pressure-flow relationship in narrow blood vessels is expressed and recommendations for further research to determine venous pressure are made for an overall mathematical model.

INTRODUCTION

Recent developments in spacecraft and high performance aircraft have resulted in the exposure of the human body to the hazards of high accelerations beyond tolerance levels. In addition to injury to various parts of the body, cardiac insufficiency and the consequent physiological malfunctions such as headache, abdominal pain, change in heart rate, chest pain, loss of vision and hemorrhage are some of the manifestations of acceleration trauma. The knowledge of the response of the cardiovascular system to acceleration stress is essential to the development of protective devices which are designed to increase acceleration tolerance by the human body during aircraft and spacecraft maneuvers. The objective of the present investigation is to understand the response of the blood vessels to transient and sustained aerospace acceleration profiles.

*Professor, Department of Engineering Mechanics

Since it is impossible to actually subject the human body to abnormal high accelerations without inflicting an injury, the response of the vascular system to transient and sustained accelerations is investigated mathematically. Theoretical analyses are extremely helpful for evaluating the relative injury potential for various acceleration functions, in guiding experimental investigations, and in developing and understanding protective measures. Mathematical procedures also provide the basis for establishing precise dynamic and physiological scaling laws needed to translate experimental data obtained with various species into meaningful results for humans.

In recent years, numerous mathematical investigations of arterial dynamics have appeared in the scientific literature. Womersley [1] and Noordergraaf [2] have presented a mathematical analysis of blood flow through arteries by using a lumped parameter model. Taylor [3], Kenner [4], Attinger et.al. [5] used distributed parameter models to analyze pressure-flow relationships in arteries and veins. Several articles related to blood flow in arteries have appeared in the book by McDonald [6]. An elastic tube theory of blood flow has been treated by Lambert [7] and Skalak and Stathis [8]. Kivity and Collins [9] presented a viscoelastic tube model for aortic rupture under decelerative forces. Rudinger [10] studied the effect of shock waves on mathematical models of aorta for better understanding of the behavior of the actual aorta.

To understand the blood flow characteristics in the arterial system, the knowledge of the material properties of the arterial wall is essential. Bergel [11], Fung [12], Demiray and Vito [13] have utilized mathematical models of the constitutive properties of the arterial tissue to determine the stresses in the arterial walls. In the present study, the strain energy function given by Demiray and Vito [13] for an arterial wall specimen has been used in determining the aortic pressure that is compatible with large deformation of the aorta and the associated flow under acceleration stress.

The present study also considers the effect of acceleration on the microcirculation. Microcirculation under normal conditions was investigated by Prothero and Burton [14], Whitmore [15], Gross and Aroesty [16], Gross and Intaglietta [17], Skalak [18] and Fung [19] who presented various theories of flow in the capillary bed connecting the arteries and veins.

Several experimental investigations on the effects of acceleration stress on the human body have been performed at the USAF School of Aerospace Medicine at Brooks Air Force Base, Texas. Burton [20] subjected miniature swine to G_z acceleration to study its effects on the organism and extrapolated the results to human beings. Parkhurst, et.al. [21] conducted experiments on human tolerance to high $+G_z$ forces. Leverett, et.al. [23] studied the cardiovascular responses during and following exposure to $+G_z$ forces in chronically instrumented anesthetized dogs. Burton and MacKenzie [24] determined the extent of heart pathology as a function of acceleration stress.

MATHEMATICAL FORMULATION

A. *Equations of Fluid Motion*

The geometry of the elastic tube containing blood in motion is shown in Fig. 1. Let r, θ, z be the cylindrical polar coordinates and let u, v, and w be the velocity components in the corresponding directions. Assuming axial symmetry in flow and tube deformation, the Navier-Stokes equations for the flow of blood can be written as:

$$\frac{\partial u}{\partial t} + u\frac{\partial u}{\partial r} + w\frac{\partial u}{\partial z} = -\frac{1}{\rho_o}\frac{\partial p}{\partial r} + \nu\left[\frac{\partial^2 u}{\partial r^2} + \frac{1}{r}\frac{\partial u}{\partial r} + \frac{\partial^2 u}{\partial z^2} - \frac{u}{r^2}\right] \tag{1}$$

$$\frac{\partial w}{\partial t} + u\frac{\partial w}{\partial r} + w\frac{\partial w}{\partial z} = -\frac{1}{\rho_o}\frac{\partial p}{\partial z} + \nu\left[\frac{\partial^2 w}{\partial r^2} + \frac{1}{r}\frac{\partial w}{\partial r} + \frac{\partial^2 w}{\partial z^2}\right] + g(t) \tag{2}$$

where p is the pressure, ν is the kinematic viscosity, ρ_o is density of blood and g(t) is the body force per unit mass caused by the acceleration. The continuity equation is

$$\frac{\partial u}{\partial r} + \frac{u}{r} + \frac{\partial w}{\partial z} = 0 \tag{3}$$

The above equations are nondimensionalized using a typical length, R_o, which is the initial (undeformed) radius of the aorta, and U, the average velocity of blood in the aorta. Introducing the new quantities

$$\begin{gathered} t^* = \frac{tU}{R_o} \quad r^* = \frac{r}{R_o} \quad z^* = \frac{z}{R_o} \quad w^* = \frac{w}{U} \\ u^* = \frac{u}{U} \quad p^* = \frac{p}{\rho_o U^2} \quad g^* = \frac{Rg}{U^2} \quad Re = \frac{UR_o}{\nu} \end{gathered} \tag{4}$$

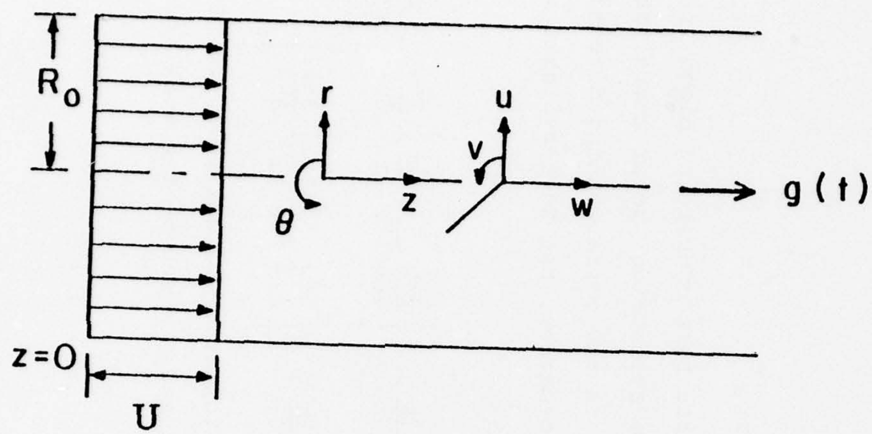


Figure 1. Definition sketch for fluid flow variables.

and deleting the "stars" for simplicity, the equations of motion and the continuity equation in terms of the newly defined variables become

$$\frac{\partial u}{\partial t} + u\frac{\partial u}{\partial r} + w\frac{\partial u}{\partial z} = -\frac{\partial p}{\partial r} + \frac{1}{Re}\left(\frac{\partial^2 u}{\partial r^2} + \frac{1}{r}\frac{\partial u}{\partial r} + \frac{\partial u^2}{\partial z^2} - \frac{u}{r^2}\right) \tag{5}$$

$$\frac{\partial w}{\partial t} + u\frac{\partial w}{\partial r} + w\frac{\partial w}{\partial z} = -\frac{\partial p}{\partial z} + \frac{1}{Re}\left(\frac{\partial^2 w}{\partial r^2} + \frac{1}{r}\frac{\partial w}{\partial r} + \frac{\partial^2 w}{\partial z^2}\right) + g(t) \tag{6}$$

$$\frac{\partial u}{\partial r} + \frac{u}{r} + \frac{\partial w}{\partial z} = 0 \tag{7}$$

The boundary and initial conditions are

$$\begin{aligned} u &= \frac{dR_1}{dt} \quad \text{at} \quad r = R_1 \qquad t \geq 0 \\ w &= 0 \quad \text{at} \quad r = R_1 \qquad t \geq 0 \\ w &= 1 \quad \text{at} \quad z = 0 \qquad t \geq 0 \end{aligned} \tag{8}$$

where R_1 is the inside radius of the blood vessel in the deformed state.

B. *Equations of Motion for Thin-walled Elastic Tube:*

The theory of large elastic deformations is utilized to describe the time-dependent deformation of the blood vessels. In view of the published results on blood pooling and the consequent cardiac insufficiency, the application of large deformation theory appears necessary. Demiray and Vito [13] have previously used this theory to calculate the deformation of arteries.

The undeformed and deformed cylindrical tubes are shown in Fig. 2. Let r, θ, z in the deformed tube. r_1, r_2 are inside and outside radii, respectively, of the undeformed tube, and R_1, R_2 those of the deformed tube. Axial stretch of the tube is neglected because of tethering caused by the surrounding tissue. Assuming the material of the blood vessels to be homogeneous, incompressible, and isotropic, the equation of motion is given by

$$\frac{\partial}{\partial R}\left[P + \beta\left(1 + \frac{1}{\lambda^2}\right) e^{\alpha(I_2 - 3)}\right] + \frac{\beta}{R}\left(\frac{1}{\lambda^2} - \lambda^2\right) e^{\alpha(I_2-3)} = \rho_w \frac{\partial^2 R}{\partial t^2} \tag{9}$$

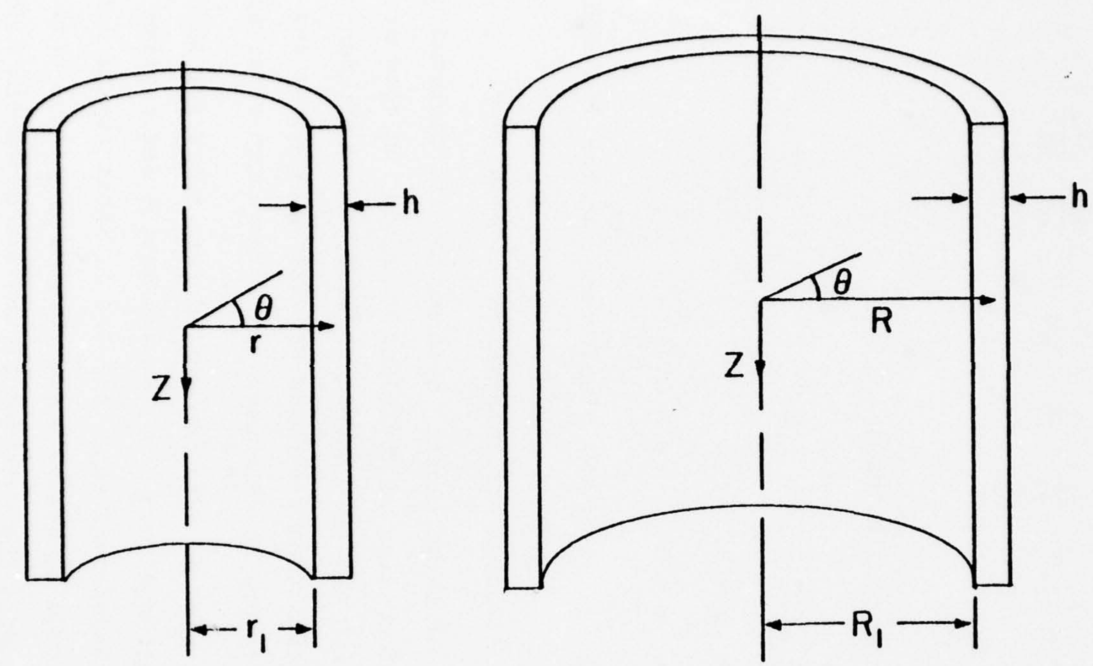


Figure 2. *Geometry of the undeformed and deformed elastic tube.*

in which α and β are material constants, $\lambda = R/r$ is the circumferential stretch ratio, and $I_2 = 1 + \lambda^2 + 1/\lambda^2$ the second strain invariant.

It has been shown that for a biomaterial, a reasonable strain energy function as shown in [13] is

$$W = \frac{\beta}{2\alpha}\left[e^{\alpha(I_2 - 3)} - 1\right] \tag{10}$$

The incompressibility condition leads to:

$$R^2 - R_1{}^2 = r^2 - r_1{}^2 \tag{11}$$

and

$$\frac{\partial^2 R}{\partial t} = -\frac{R_1{}^2}{R^3}\left(\frac{dR_1}{dt}\right)^2 + \frac{1}{R}\left(\frac{dR_1}{dt}\right)^2 + \frac{R_1}{R}\,\frac{d^2R_1}{dt^2} \tag{12}$$

With p_1, p_2 denoting the pressure on the inside and outside wall, respectively, of the blood vessel, the boundary conditions become

$$\begin{aligned} \tau^{11} &= -p_1(t) \text{ at } R = R_1 \\ \tau^{11} &= -p_2(t) \text{ at } R = R_2 \end{aligned} \tag{13}$$

Introduction of Eqs. (12) and (13) into Eq. (9), and using the dimensionless variables

$$P^* = \frac{P}{\rho_o U^2}\,,\quad R_1{}^* = \frac{R_1}{R_o}\,,\quad R_2{}^* = \frac{R_2}{R_o}\,,\quad t^* = \frac{tU}{R_o} \quad \beta^* = \frac{\beta}{\rho_o U^2} \quad \rho_w{}^* = \frac{\rho w}{\rho_o} \tag{14}$$

yield after integration and deletion of 'stars' for simplicity the equation of motion in the form

$$p_1(t) - p_2(t) = \rho_w R_1 \frac{d^2R_1}{dt^2}\ln\frac{R_2}{R_1} - \rho_w\left(\frac{dR_1}{dt}\right)^2\left[\ln\frac{R_2}{R_1} + \frac{1}{2}\left(\frac{R_2{}^2}{R_1{}^2} - 1\right)\right]$$

$$- \beta\int_{\lambda_1}^{\lambda_2} \frac{1+\lambda^2}{\lambda^3}\, e^{\alpha(\lambda^2 + 1/\lambda^2 - 2)}\, d\lambda \tag{15}$$

The initial conditions are:

At time $t = t_o$, $R_1 = R_o$, $dR_1/dt = u$, radial velocity of fluid.

C. *Effect of Acceleration on Microcirculation*

The blood vessels of microcirculation are extraordinarily small, and their typical dimensions are of the order of microns. Under normal circumstances, the velocity of the blood in the microcirculation is 1 mm/sec and the Reynolds number is of the order $O(10^{-3})$, which is sufficiently small so that the Stokes flow approximations are applicable. Neglecting the inertial effects and assuming that the stream lines are nearly parallel, the dimensionless equation of fluid motion is the axial (z) direction becomes

$$\frac{\partial w}{\partial t} = - \frac{\partial p}{\partial z} + \frac{1}{Re}\left(\frac{\partial^2 w}{\partial r^2} + \frac{1}{r}\frac{\partial w}{\partial r} + \frac{\partial^2 w}{\partial z^2}\right) + g(t) \qquad (16)$$

which can be rearranged to read

$$Re \frac{\partial w}{\partial t} = - Re\left(\frac{\partial p}{\partial z}\right) + \left(\frac{\partial^2 w}{\partial r^2} + \frac{1}{r}\frac{\partial w}{\partial r} + \frac{\partial^2 w}{\partial z^2}\right) + Re\ g(t) \qquad (17)$$

In the earth's natural gravitational field, the dimensionless g, as given in Eq. (4), is of the order $O(10^{-2})$, and with the effect of $Re \sim O(10^{-3})$ in the last term $Re\ g(t)$ in Eq. (17) becomes physiologically insignificant, being of the order $O(10^{-5})$. We estimate that the effect of acceleration on microcirculation *per se* can be safely neglected up to 100 g. However, the pressure of the blood pooled in the arteries and veins can affect the flow rate in the small vessels. For this reason it is necessary to determine a relationship between the pressure gradient and the flow rate in the small blood vessels.

For the flow of a Newtonian fluid in a uniform tube Szymanski [27] showed that the flow would be fully developed if $\nu t/D^2 \gtrsim 1$, where t = time, ν = kinematic viscosity, and D = tube diameter. An extension of this criterion to microcirculation yields $\nu\Delta t/D^2 \gtrsim 1$ for flow to be quasi-steady, where Δt is the smallest characteristic time of the unsteadiness in flow. According to Burton [28], $\Delta t \simeq 0.1$ sec; using $\nu = 0.04$ Stokes, one finds that the diameter D must be greater than 600μ (microns) for any significant changes in flow due to unsteadiness become entirely negligible. On this basis Benis [29] argued that the effect of unsteadiness on non-Newtonian flow could also be neglected. Thus, the use of steady-flow equations can be justified for microcirculation.

For steady capillary flow, the flow rate through a circular tube can be expressed by

$$Q = 2\pi \int_{o}^{R} rw \, dr \tag{18}$$

where Q = flow rate, R = tube radius, and w = blood velocity. Integration by parts of the right hand side yields

$$Q = \pi \int_{o}^{R} d(r^2 w) - \pi \int_{o}^{R} r^2 \left(\frac{dw}{dr}\right) dr \tag{19}$$

The first integral on the right hand side of Eq. (19) is zero. In the second integral the domain of integration can be divided into two regions: a cone of unsheared fluid extending to radius R_y, and the annular region bounded by the unsheared fluid and the tube wall.

Then,

$$Q = -\pi \int_{o}^{R_y} r^2 \left(\frac{dw}{dr}\right) dr - \pi \int_{R_y}^{R} r^2 \left(\frac{dw}{dr}\right) dr \tag{20}$$

The first term on the right hand side in Eq. (20) which represents the core integral is zero. By changing the variables from r to τ in the second term as suggested by Merrill et.al. [30], Eq. (20) can be written as

$$Q = \frac{8\pi}{(\Delta P/L)^3} \int_{\tau_y}^{\tau_w} \tau^2 \dot{\gamma} \, d\tau \tag{21}$$

where L = length of the capillary

ΔP = pressure drop

τ = shear stress

$\dot{\gamma}$ = shear rate.

The shear stress and the shear rate are related by an empirical equation

$$\tau^{1/2} = \tau_y^{1/2} + \mu^{1/2} \dot{\gamma}^{1/2} \tag{22}$$

in which τ_y is the yield shear stress, and $\mu^{1/2}$ is a constant which represents the slope of the Casson plot relating the viscometric parameters of blood. In Poiseuille's flow μ becomes the blood viscosity. Substitution of Eq. (22) into Eq. (21) and integration yields

$$Q = \frac{\pi R^4 (\Delta P/L)}{8\mu} - \frac{4\pi \tau_y^{1/2} R^{7/2} (\Delta P/L)^{1/2}}{7\sqrt{2}\,\mu} - \frac{2\pi \tau_y^4}{21(\Delta P/L)^3 \mu} + \frac{\pi \tau_y R^3}{3\mu} , \tag{23}$$

which is valid under the assumption that the flow is steady, laminar and incompressible, and blood is homogeneous. In the above equation, τ_y and μ are known constants; then plots of Q vs. $\Delta P/L$ for capillaries of various radii can be easily constructed. An example of this relationship is shown in Fig. 6.

RESULTS AND DISCUSSION

To determine the response of the blood vessels one must solve Eqs. (5), (6), (7) and (15) under the appropriate boundary conditions. Since these equations were coupled a numerical solution using a digital computer was sought as described in an earlier paper [31].

The following constants were used in the solution: $R_o = 1.47$ cm, $\nu = 11.9$ cm/s, $\rho_w = 1.05$ gr/cm^3, $\rho_o = 1.05$ gr/cm^3, $\nu = 0.038$ Stoke, $\alpha = 0.8$ and $\beta = 11.35 \times 10^4$ dynes/cm^2. The constants α and β are elastic constants which appear in the strain energy function W, Eq. (10). These values are reported in literature for a specimen of human aorta. The elastic constants for specimens of veins are not available, primarily because the research of determining the strain energy function is the form of Eq. (10) is relatively new.

As examples, two deceleration profiles were used in the solution: one, a linear, monotonically increasing function of time represented by $g(t) = 7840t$ cm/s^2, and a transient type (Fig. 3) which increased and decreased rapidly [32]. For both these cases, aortic pressures were calculated by utilizing a finite difference technique which involved Runge-Kutta integration procedure and Adams-Bashforth predictor-corrector method. The dimensions of the aorta were chosen from the physiological data presented in Westerhof

et.al. [33]. The aortic pressures calculated for the sustaining linear deceleration profile is shown in Fig. 4. The aortic pressure in response to the transient profile of Fig. 3 is presented in Fig. 5 along with an experimentally determined pressure in the thoracic aorta of a beagle dog which was subjected to the same transient deceleration profile for comparison. The shapes of the pressure vs. time curves are nearly the same indicating satisfactory qualitative agreement. An exact quantitative agreement cannot be expected because of the anatomical and physical differences between the subjects.

The method of calculation described above for an artery can be easily extended to a vein, and the pressure difference across the capillary bed can be determined. Equation (23) represents a relationship between the flow rate and the pressure gradient in a narrow tube. For selected values of $\tau_y = 0.042$ dyne/cm^2, $\mu = 0.05$ dyne-sec/cm^2, D = 100 microns, L = 2.5 cm, the flow rate Q is computed for various pressure drops in the range $\Delta P/L = 10 \sim 10000$ dyne/cm^3 and shown in Fig. 6. If the pressure of the blood pooled on the venous side is known, the effect of acceleration on the flow in microcirculation can be determined from Fig. 6. It must be noted that the flow rate indicated in Fig. 6 is for a narrow blood vessel of specific dimensions. To obtain the total blood flow one must formulate the solution on a statistical basis which includes the arterioles, venules and capillaries of various dimensions and changing rheological properties of blood.

The results presented in this paper are part of an effort to describe the cardiovascular system in terms of its physical and mechanical properties. Most investigations hither to reported have dealt with electrical analogs of the cardiovascular system in which various parameters were introduced in terms of resistances, impedances and capacitances. These electrical quantities may not truly represent the cardiovascular parameters under high g conditions, and an analysis of the cardiovascular system which is based on its original properties is therefore desirable.

RECOMMENDATIONS FOR FURTHER RESEARCH

The model presented in this report represents the behavior of major arteries and small blood vessels. The method which is used to model the behavior of the arteries can also be adapted to describe the response of the veins by introducing the appropriate material constants and dimensions of the

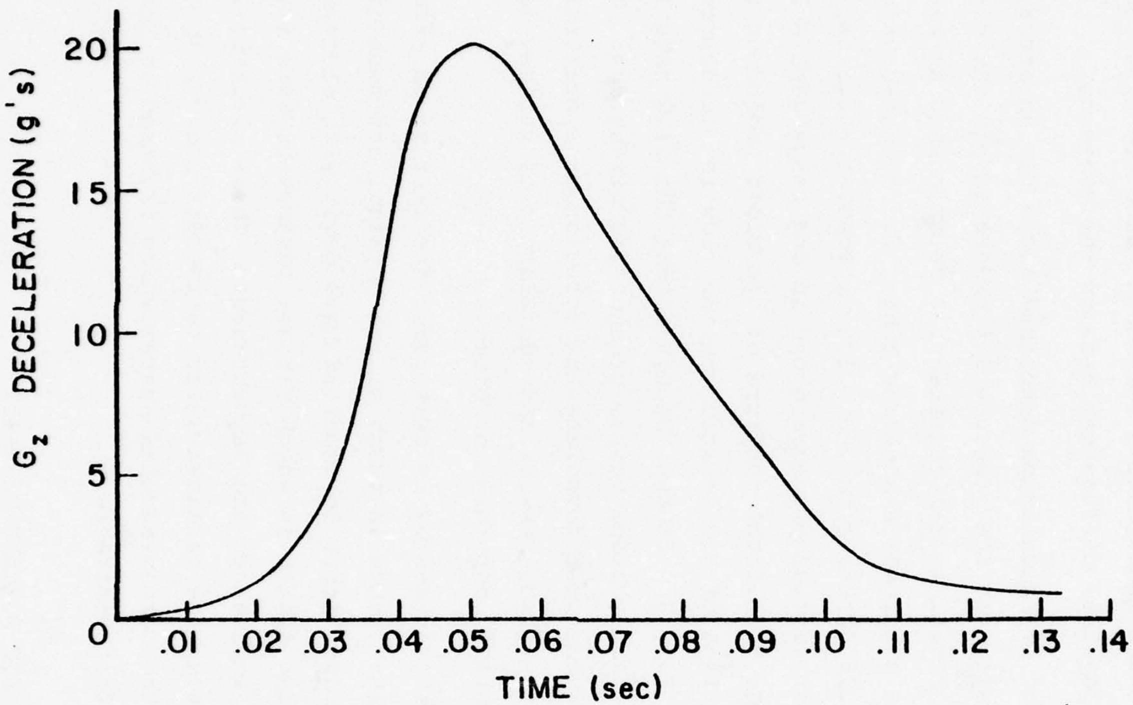


Figure 3. Deceleration function measured in the experiment by Hanson [32].

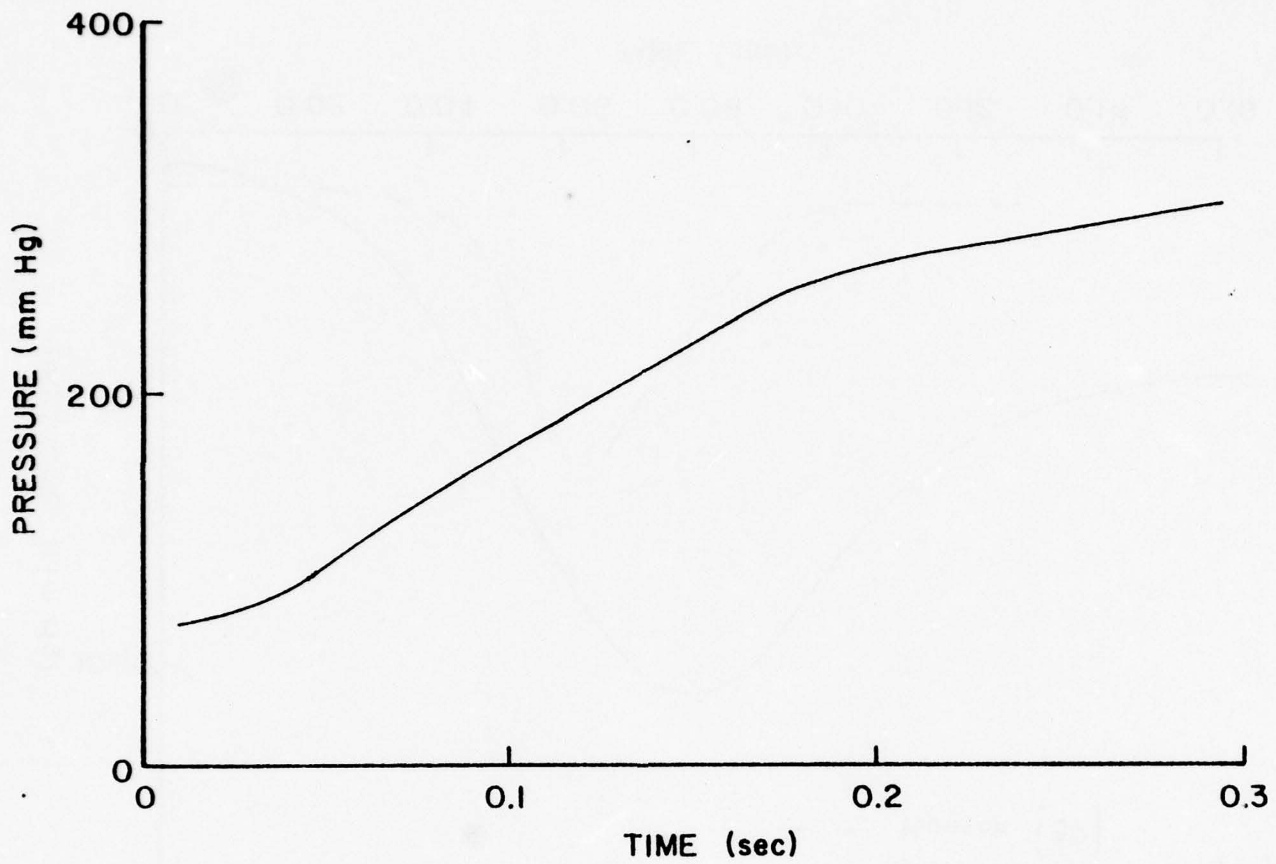


Figure 4. Calculated pressure in the aorta for the linear deceleration $g(t) = 7840\ t\ cm/s^2$.

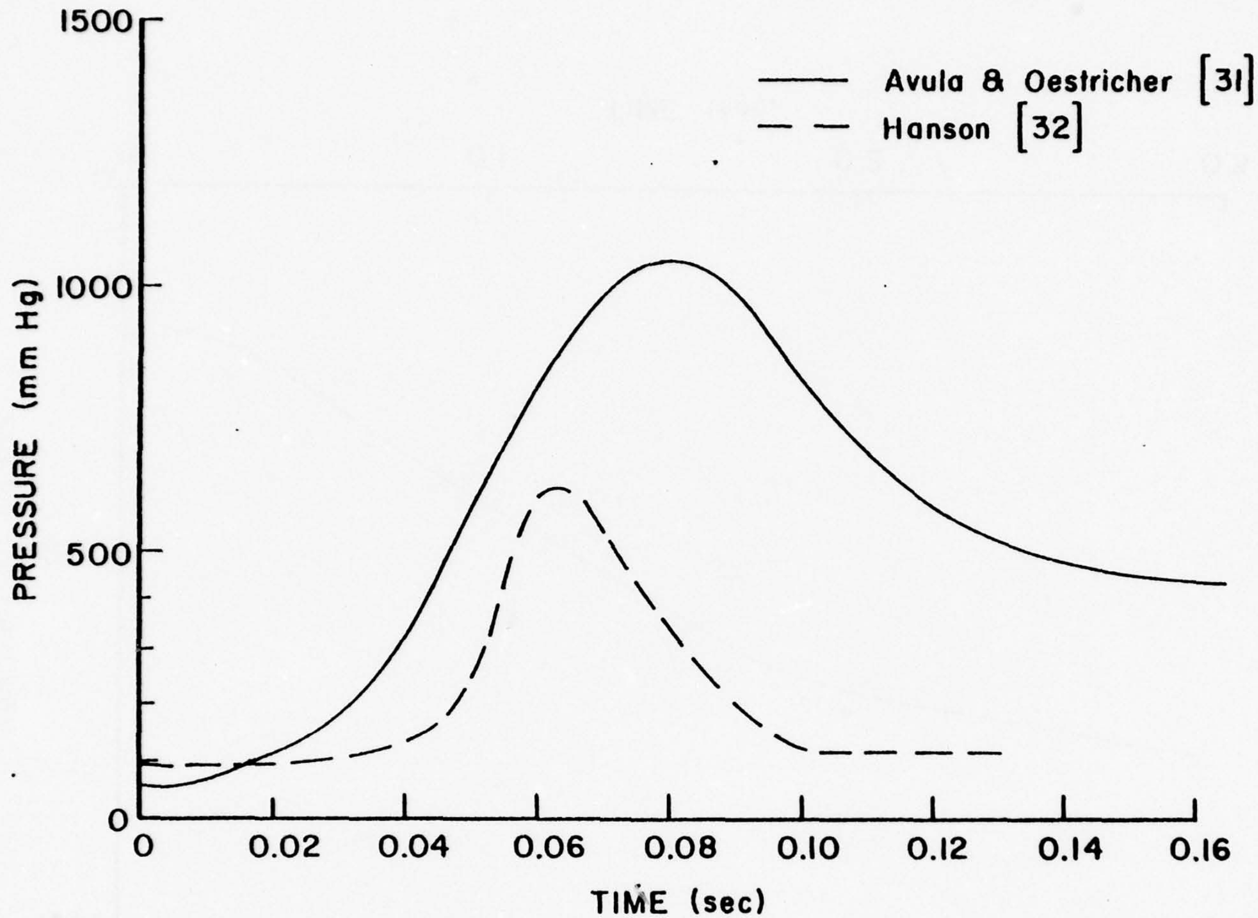


Figure 5. Comparison of theoretical and experimental pressures for the deceleration in Hanson's [32] experiment.

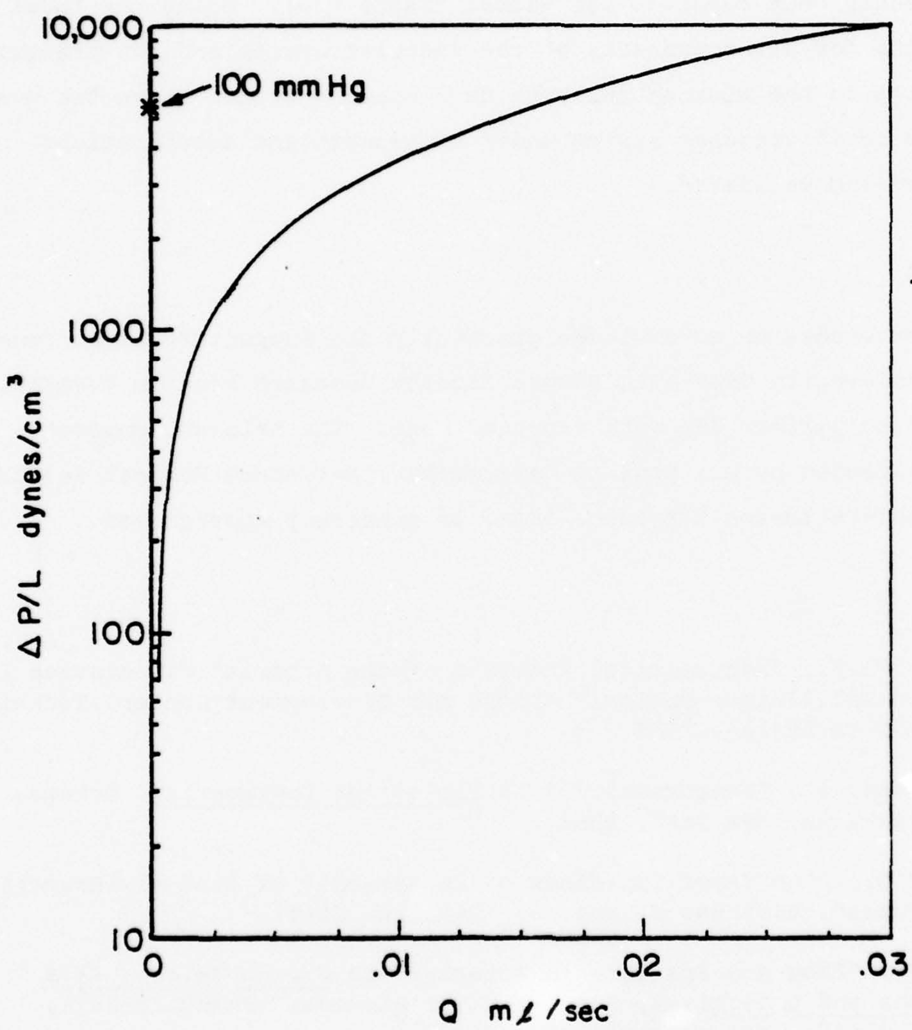


Figure 6. Pressure - Flow relationship for a narrow blood vessel (Eq. 33).

veins. At present the constants α and β for the material of the venous wall are not available in the literature. Investigations to determine these constants for veins must be undertaken. Since α and β vary from location to location they must be determined throughout the vascular network.

The pressure flow relationship in the human microvasculature must be developed on a statistical basis taking into account the branching patterns and the resultant hierarchy of vessels between the arteries and veins. Such a study has recently been reported for animal tissue [34]. Using the local material constants for all components of the vascular system and the pressure-flow relationships in the microvasculature developed as stated above the overall model of the cardiovascular system under time-dependent accelerations must be generated and validated.

ACKNOWLEDGEMENTS

The author wishes to acknowledge gratefully the support received from The Ohio State University USAF-ASEE Summer Faculty Research Program (WPAFB), June-August 1978 to perform the work reported here. The help and constant encouragement extended by Dr. Hans L. Oestreicher, Aerospace Medical Research Laboratory, Wright-Patterson Air Force Base, is sincerely appreciated.

REFERENCES

[1] Womersley, J.R., "Mathematical Analysis of the Arterial Circulation in a State of Oscillatory Motion," Wright Air Development Center, Technical Report WADC-TR-56-164, 1958.

[2] Noordergraaf, A., "Hemodynamics," in Biological Engineering, Schwan, H., ed., McGraw-Hill, New York, 1969.

[3] Taylor, M.D., "The Input Impedance of an Assembly of Randomly Branching Elastic Tubes", Biophys. J. Vol. 6, 1966, pp. 25-51.

[4] Kenner, T., "Flow and Pressure in Arteries" in Biomechanics: It's Foundations and Objectives, Fung, Y.C. et al. eds., Prentice-Hall, Englewood Cliffs, New Jersey, 1972.

[5] Attinger, E.O., Anne, T., Mikami, T., and Sugawara, H., "Modeling of Pressure Flow Relationships in Veins". Hemorheology, Copley, A., ed., McGraw-Hill, New York, 1964.

[6] McDonald, D.A., Blood Flow in Arteries. E. Arnold (Publishers) Ltd., London, 1960.

[7] Lambert, J.W., "On the Nonlinearities of Fluid Flow in Nonrigid Tubes", J. Franklin Inst., Vol. 26, 1958, pp. 83-102.

[8] Skalak, R. and Stathis, T., "A Porous Tapered Elastic Tube Model of a Vascular Bed", in Biomechanics, Fung, Y.C., ed., ASME Symposium, New York, New York, 1966.

[9] Kivity, Y. and Collins, R., "Nonlinear Wave Propagation in Viscoelastic Tubes: Application to Aortic Rupture", J. Biomechanics, Vol. 7, 1974, pp. 67-76.

[10] Rudinger, G., "Shock Waves in Mathematical Model of the Aorta", J. Appl. Mech. ASME Vol. 37, 1970, pp. 34-37.

[11] Bergel, D.H., "The Dynamic Elastic Properties of the Arterial Wall", J. Physiol. Vol. 156, 1961, pp. 458-469.

[12] Fung, Y.C., "Elasticity of Soft Tissues in Simple Elongation", Am. J. Physiol. Vol. 213, 1967, pp. 1532-1544.

[13] Demiray, H., and Vito, R.P., "Large Deformation Analysis of Soft Biomaterials", Developments in Theoretical and Applied Mechanics, 8, Proc. 8th Southeast. Conf. on Theor. and Appl. Mech., 1976, pp. 515-522.

[14] Prothero, J. and Burton, A.C., "The Physics of Blood Flow in Capillaries. I. The Nature of the Motion", Biophys. J. Vol. 1, 1961, pp. 565-579.

[15] Whitmore, R.L. "A Theory of Blood Flow in Small Vessels", J. Appl. Physiol. Vol. 22, No. 4, 1967, pp. 767-771.

[16] Gross, J.F. and Aroesty, J., "Mathematical Models of Capillary Flow: A Critical Review" Biorheology, Vol. 9, 1972, pp. 225-264.

[17] Gross, J.F. and Intaglietta, M., "Effects of Morphology and Structural Properties on Microvascular Hemodynamics" Report No. P-5000, The Rand Corporation, Santa Monica, California, 1973.

[18] Skalak, R., "Mechanics of the Microcirculation", in Biomechanics: It's Foundations and Objectives, Fung, Y.C. et. al. eds., Prentice-Hall, Englewood Cliffs, New Jersey, 1972.

[19] Fung, Y.C., "Microscopic Blood Vessels in the Mesentery", Biomechanics, Fung, Y.C., ed., ASME Symposium, New York, New York, 1966.

[20] Burton, R.R., "Positive ($+G_z$) Acceleration Tolerances of the Miniature Swine: Application as a Human Analog", Aerosp. Med., Vol. 44, 1973, pp. 294-298.

[21] Parkhurst, M.J., Leverett, S.D., Jr., and Shurbrooks, S.J., Jr., "Human Tolerance to High, Sustained $+G_z$ Acceleration", Aerosp. Med., Vol. 43, 1972, pp. 708-712.

[22] Leverett, S.D., Jr., Burton, R.R., Crossley, R.J., Michaelson, E.D., and Shurbrooks, S.J., Jr., "Physiologic Responses to High Sustained $+G_z$ Acceleration", USAF School of Aerospace Medicine, Tech. Rep. No. 73-21, 1973.

[23] Peterson, D.F., Bishop, V.S., and Erickson, H.H., "Cardiovascular Changes During and Following 1-min Exposure to $+G_z$ Stress", Aviat. Space Environ. Med., Vol. 46, No. 6, 1975, pp. 775-779.

[24] Burton, R.R., and MacKenzie, W.F., "Heart Pathology Associated with Exposure to High Sustained G_z", Aviat. Space Environ. Med., Vol. 46, No. 10, 1975, pp. 1251-1253.

[25] Green, A.E. and Zerna, W., Theoretical Elasticity, Oxford University Press, New York, London, 1968.

[26] Klip, W., The Foundations of Medical Physics, University of Alabama Press, Birmingham, Alabama, 1969.

[27] Ronse, H. (ed.) Advanced Mechanics of Fluids, Wiley, New York, 1959.

[28] Burton, A.C., Physiology and Biophysics of the Circulation, Year Book Medical Publishers, Chicago, 1965.

[29] Benis, A.M., "Laminar Flow of Power-law Fluids Through Narrow Three-dimensional Channels of Varying Gap. Chem. Engr. Sci., Vol. 22, 1967, pp. 805-822.

[30] Merrill, E.W., Benis, A.M., Gilliland, E.R., Sherwood, T.K., and Salzman, E.W., "Pressure-Flow Relations of Human Blood in Hollow Fibres at Low Flow Rates", J. Appl. Physiol. Vol. 20, No. 5, 1965, pp. 954-967.

[31] Avula, X.J.R. and Oestreicher, H.L., "Mathematical Model of the Cardiovascular System Under Acceleration Stress", Aviat. Space Environ. Med., Vol. 49, No. 1, 1978, pp. 279-286.

[32] Hanson, P.G., "Pressure Dynamics in Thoracic Aorta During Linear Deceleration", J. Appl. Physiol. Vol. 28, No. 1, 1970, pp. 23-27.

[33] Westerhof, N., Bosman, F., Devries, C.J., and Noordergraaf, A., "Analog Studies of the Human Systemic Arterial Tree", J. Biomech. Vol. 2, 1969, pp. 121-143.

[34] Zweifach, B.W. and Lipowsky, H.H., "Quantitative Studies of Microcirculatory Structure and Function - III. Microvascular Hemodynamics of Cat Mesentery and Rabbit Omentum", Circ. Res. Vol. 41, No. 3, 1977, pp. 380-390.

EVOKED RESPONSE MEASURES OF RESOURCE ALLOCATION:

EFFECTS OF VARYING THE PRIMARY TASK WORKLOAD

Lloyd F. Elfner* and Robert D. O'Donnell**

Florida State University, Tallahassee, Florida

Abstract

The amplitude of the P_3 component of the evoked response potential has been shown to reflect active processing of information. The present study assessed the value of P_3 in providing an index of the resource demands of a forced function tracking task. The P_3 component was generated by a concomitant secondary task in which a silent count was kept of the rare occurring auditory tone presented in a random Bernoulli sequence. The two tones were presented with a random inter-stimulus interval. Past research has shown the P_3 amplitude to drop precipitously when a primary tracking task is involved. Preliminary measurements of the obtained P_3 amplitudes indicated notable P_3 even under a difficult primary tracking task. A significant interaction between tracking task load and probability of occurrence of the rare tone was found and the P_3 amplitude was shown to vary systematically with probability of the rare tone. The latter result has been replicated many times thus indicating the validity of the experimental procedures. The above findings suggest parametric research to delineate factors in the tracking task that produce systematic changes in P_3 amplitude should be undertaken.

* Professor, Department of Psychology

** Lt Col, United States Air Force

EVOKED RESPONSE MEASURES OF RESOURCE ALLOCATION: EFFECTS OF VARYING THE PRIMARY TASK WORKLOAD

INTRODUCTION

The present study is a replication of a portion of preliminary experimentation reported by Donchin (1976). The major thrust of the research is to assess the utility of employing event-related potentials (ERP) in a procedure that simulates to some degree the control and monitoring tasks of a pilot. The assumption is that the ERP reflects the subject's information processing characteristics. The procedure used is the "work manipulation paradigm". In this paradigm, probes are presented along a secondary channel that is unrelated to a primary manual tracking task. Obviously there is a limited ability to perform in terms of processing incoming information. A direct result of this assumption is that if one extends the resources applied to a primary task there will be a decrement in secondary task performance. The "residual" attention is assessed in an uncertainty task in which ERPs are measured to the probe stimuli under different levels of primary task load. The use of the uncertainty principle is relatively recent in psychophysiological research.

Since the germinal study by Sutton et al(1965), much work has been done in the area of stimulus uncertainty and its effect on the evoked brain potentials measured in humans. Sutton et al (1965) suggested that uncertainty is reflected by two kinds of influences. The first is largely exogenous and probably is related to characteristics of the stimulus. The second is endogenous and related to reaction or attitude of the subject.

The exogenous influence is reflected in the early components of the event related potentials and the endogenous influence is reflected in the late components of the evoked response. Both early and late components appear to change systematically with change in stimulus uncertainty in that the N_1 component and the P_3 component are considerably enhanced when stimulus uncertainty is present. Due to the fact that the P_3 component is considered to reflect subjective influences, the cognitive structure underlying the P_3 endogenous component has been studied extensively in an effort to understand the "process" that underlies the reflected changes. In fact, it has been suggested (Donchin, Ritter, and Mc Callum, 1977) that the process associated with the endogenous event related potential components are the same processes inferred by cognitive psychologists.

The prospects for uncovering information processing abilities through the use of event related potentials has attracted numerous investigators to the field. Much work has been done on probability and P_3 (Corby and Kopell, 1973; Donchin, Kubovy, Kutas, Johnson, and Herning, 1973; Friedman, Hakerem, Sutton, and Fleiss, 1973; Rohrbaugh, Donchin, and Ericksen, 1974; Ruchkin, Sutton, and Tueting, 1975; Tueting and Sutton, 1973; Tueting, Sutton, and Zubin, 1971).

Among the several paradigms for stimulus presentation developed to study the uncertainty effect is the oddball technique in which a sequence of two or more stimuli is presented in a random manner and the subject is instructed to internally count the target stimulus. Typically the amplitude of the P_3 component shows a decrease with an increase in probability of the occurrence of the target stimulus (Courchesne, Hillyard, and Galambos, 1975; Hillyard, Hink, Schwent, and Picton, 1973; Mast, and Watson, 1968; Picton and Hillyard, 1974; Ritter, and Vaughn, 1969; Roth, Ford, Lewis, and Kopell, 1976; Squires, Donchin, Herning, and McCarthy, 1977; Squires, Donchin, Squires, and Grossberg, 1977). Also, Duncan-Johnson and Donchin, 1977 have demonstrated that the probability-amplitude relationship holds across all probabilities of target occurrence.

Most of the above studies have employed regular inter-stimulus intervals in the signal presentations. However, two studies appear relevant to the temporal structure effects of the stimulus presentation. Ford, et al (1976) demonstrated that uncertainty in time of presentation might have a stronger effect on P_3 than sequential probability. However, McCarthy and Donchin (1976) found that temporal uncertainty reduced the P_3 amplitude as opposed to a situation in which the subject dictated the temporal sequence of stimulus presentation. The question of the effect on temporal uncertainty effects on P_3 remains problematical.

The fact than N_1 is clearly more than an indicant of sensory input was shown by Picton, et al (1971). The reader will recall that Sutton (1965) demonstrated increases in amplitude of the N_1 component to uncertain stimuli. However, Picton demonstrated that in an auditory attention task, changes were evident in the N_1 component while the auditory nerve showed no change

as a function of the task relevance of the stimuli. Although exogenous in nature the N_1 component can be under control of the endogenous structures of the cortex (Donchin, et al, 1977).

Hillyard (1977) proposed a model in which the N_1 component can be considered an element of the short term memory portion of the information processor while P_3 represents a more complex contextual updating system independent of stimulus channel. The N_1 response varies as a function of the processing accorded the stimulus. The vertex centered N_1 wave is the earliest evoked response component that has been related to selective attention in man and together with its anterior scalp distribution and non-specificity with respect to modality of stimulus is suggestive of a general attentional system that is situated at a rather high cortical level.

Following the rationale developed by Hillyard in a task where relevant and non-relevant stimulus are presented in random sequence, one would predict differences in event related potentials in the N_1 component between the two types of stimuli under easily analyzable conditions. The model would predict a significant difference in the P_3 component given task relevance of one of the stimulus types.

Clearly, the paradigm employing a primary manual tracking task and concurrent secondary auditory detection task is not completely relevant to selective attention. However, one might pose the same question raised in selective attention. That is, how does the subject distribute his resources between the two tasks?

From results of past research (Israel, Wickens, and Donchin, 1978; Wickens, Israel, and Donchin, 1977) one would predict that P_3 should show a decrement in amplitude on the secondary task with primary task difficulty increase. If one assumes the hierarchally ordered stages of processing

hypothesis defined by Hillyard (1977), one might predict the N_1 component to the target stimuli will be less affected by changes in primary task load than the P3 component given that the frequency difference of the secondary task is easily analyzable.

METHOD

Subjects: Six paid subjects with normal hearing and vision were employed in the study. Ages of subjects ranged from 16 to 29 years. All subjects were right handed.

Apparatus: A computer controlled cathode ray tube display provided the moving cursor and stationary target. A forcing function method was used to drive the stimulus from the center position and the rate of slew was controlled by a ten-turn pot. A lightly loaded lever was used to provide control of the cursor. A micro-computer (Intel 80/20) was used to control auditory stimulus intervals and to provide markers for signals to be averaged. The auditory stimuli were produced by computer triggering a pair of signal generators. The signals were 1.0 KHZ and 1.3 KHZ with 5 msec zero crossing rise time 25 msec steady state and 10 msec decay time. The tones were presented binaurally at 72 db SPL through earphones (Maico, Model TDH-39) and delivered to a sound proof room (IAC, Model 600) where the subject listened.

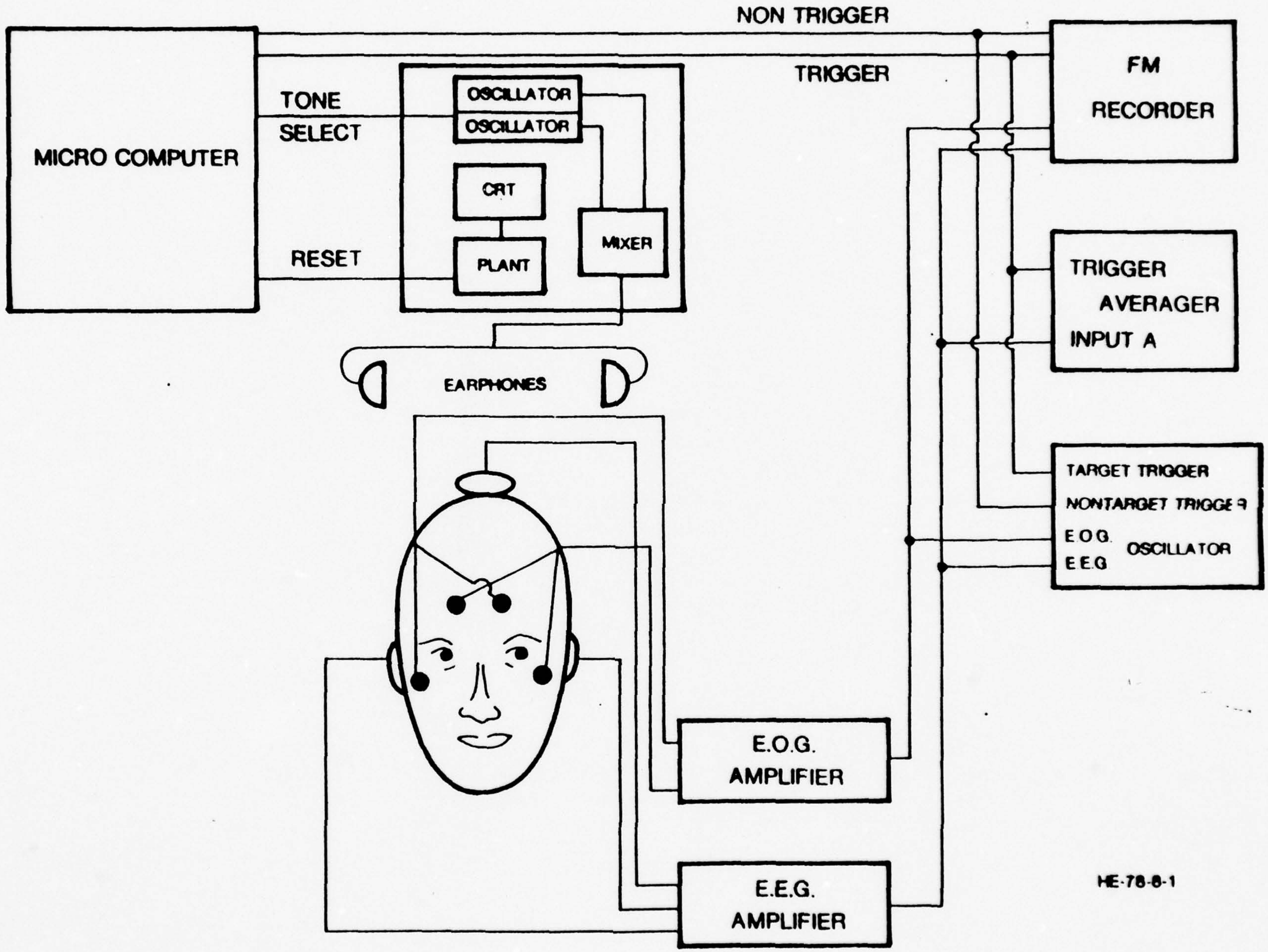
EEG pick up was accomplished by Beckman Ag Ag/Cl electrodes. EEG electrodes were placed on the vertex (active), the right mastoid (ground) and the left mastoid (reference). EOG electrodes were placed on the

superviliary ridge and just temporal to the external angular process of the contralateral orbit. Leads from these sites were shorted and the summated activity was led to the active input of the amplifier. The system was used in a symmetric fashion and the opposing lead was led to the referent input of the amplifier. A common ground was used for all measurements. The signals were fed through high impedance probes into amplifiers (Grass, Model P511) with half amplitude bandpass of 0.1 and 100HZ and amplification of 10^5. A polygraph (Grass, Model 6) was used to provide a permanent copy of the EEG, EOG, and event markers. The signals were also fed to an FM tape recorder (Honeywell, Model 5600) for later averaging. An averager (Nicolet, Model 1072) was used to provide immediate copy of summated potentials to the target stimuli and the concurrent EOG. See Figure 1 for block diagram of the apparatus.

<u>Procedure</u>: Initially, subjects were tested for vision and hearing function. Instructions were given on the procedures of the study. The subject was given an hour of practice on the manual tracking task. Six ten-minute sessions were held. Each ten minute session was broken down into one minute epochs. An adaptive procedure was employed and two criteria were used. If the subject had a small error rate at a specific slew rate the forcing function was increased on the following epoch. Gradually the slew rate was increased until the cursor was lost (cursor drifted off edge of screen) at least once during 9 of the 10 epochs. The slew rate at this position was defined as the most difficult rate for the subject for the experiment. The easy slew rate was defined as the last session where the subject did not lose the cursor. The intermediate difficult level was defined as the slew rate where the subject lost the

FIGURE 1. BLOCK DIAGRAM OF APPARATUS

BLOCK DIAGRAM OF APPARATUS



HE-78-8-1

cursor on two epochs of a session.

A second hourly practice session was held during which the subject tracked at the three levels of difficulty for periods approximating the experimental sessions. During this session the subject was also given concomitant training in the oddball task in which a silent count of the target auditory stimulus was maintained. Two sessions for each level of probability of occurrence of the target stimuli were used (10%, 20%, and 30%). Half the subjects received the 1.3 KHZ signal and half received the 1.0 KHZ signal as the target.

The Bernoulli sequences of targets and non-targets were presented in random order with random inter-stimulus interval (ISI). The average ISI was 1.75 seconds. ISIs of less than 400 msec or more than 4000 msec were excluded from the sequence. Also averaged targets and non-targets had to be preceded by an ISI of at least 850 msec. The ratio of averaged targets to total targets was held approximately equal to the ratio of averaged non-targets to total non-targets. Sixteen targets were averaged in each of the a-priori probability conditions. The number of averaged non-targets varied according to a-priori probability. The durations of the runs for the three a-priori probabilities (10%, 20%, 30%) were 11 min., 5 1/2 min., 3 1/2 min. respectively.

Experimental data was obtained in two hourly sessions. Each subject served in each of ten conditions; a control run for the 10% target probability occurrence during which the subject did not track but only counted the auditory targets and the nine conditions composed of the threemanual tracking difficulty levels and the three target probabilities in the oddball task. Conditions were randomly presented. A repeated measures analysis of variance was used to analyze the obtained N_1 and P_3 amplitudes obtained from the averages. See appendix 1 for instructions read to subjects.

RESULTS

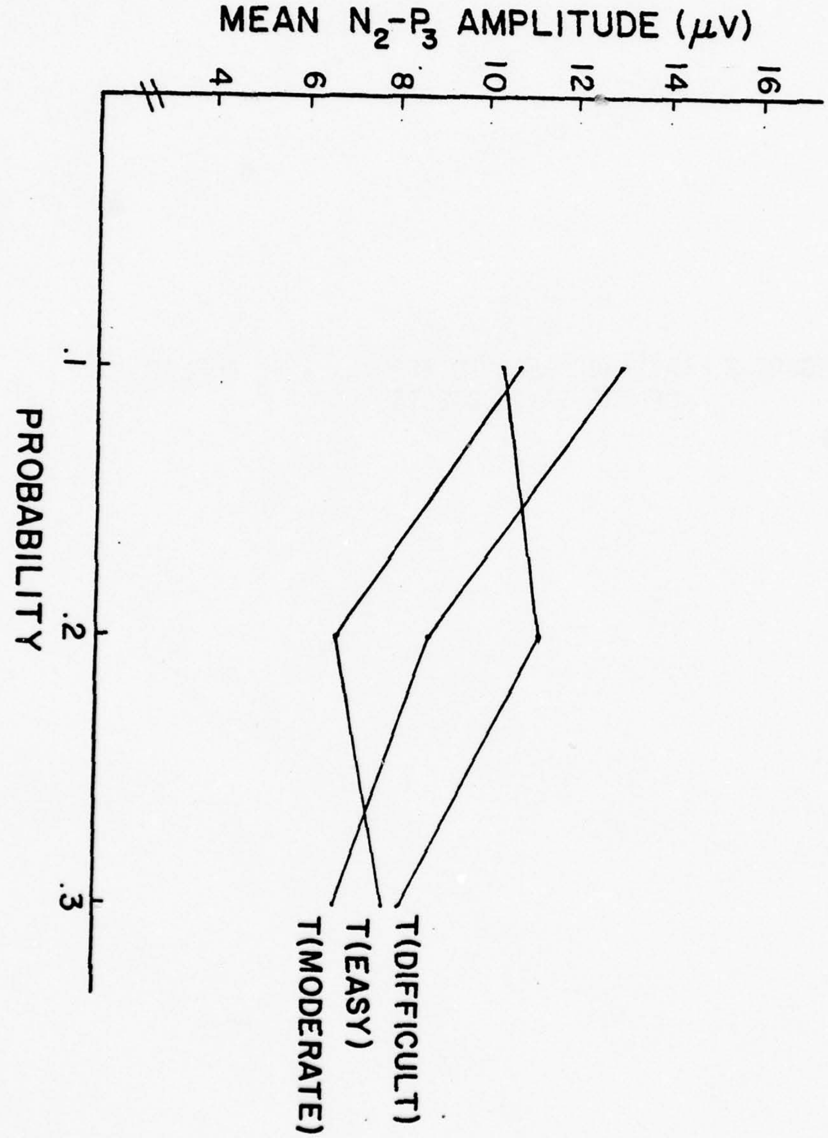
A preliminary analysis on the EEG was performed on the 16 averaged targets for the 6 subjects. Peak to peak amplitude was measured for the N_2 - P_3 components to obtain a measure of the P_3 component and a similar procedure was used on the N_1 - P_2 peak to peak amplitude to obtain a measure of the N_1 component. Table I shows the results of an analysis of variance performed on the N_2 - P_3 amplitudes. The a-priori

Table I Results of Analysis of Variance and the N_2 - P_3 Amplitudes to Target Stimuli

Source	df	SS	F	P
Probability (P)	2	128.04	12.87	.002
Task difficulty (T)	2	20.59	.69	.53
P x T	4	66.96	2.92	.046
S	5	470.81		
S x P	10	49.74		
S x T	10	148.49		
S x P x T	20	114.59		
Total	53	999.25		

probability is significant at the p=.01 level of confidence and the task difficulty by a-priori probability interaction is significant at the p=.05 level of confidence. A graph of mean values for the 6 subjects for the N_2 - P_3 amplitudes is shown in Figure 2. Table II shows the results of an analysis of variance performed on the N_1 - P_2 amplitudes.

FIGURE 2. AVERAGE PEAK TO PEAK N_2 - P_3 AMPLITUDES OF THE SIX SUBJECTS


MEAN N_2-P_3 AMPLITUDE (μV)
4
6
8
10
12
14
16
PROBABILITY
.1
.2
.3
T(DIFFICULT)
T(EASY)
T(MODERATE)

The only significant effect is the a-priori probability of the target stimulus ($p < .01$).

Table II Results of Analysis of Variance on $N_1 - P_2$ Amplitudes to Target Stimuli

Source	df	SS	F	P
Probability (P)	2	213.59	8.29	.008
Task difficulty (T)	2	11.59	.40	.68
P x T	4	33.41	.54	.71
S	5	502.31		
S x P	10	128.85		
S x T	10	143.52		
S x P x T	20	305.48		
Total	53	1338.75		

A graph showing the mean values for the six subjects for the $N_1 - P_2$ amplitudes is shown in Figure 3. Two subjects demonstrated larger $N_2 - P_3$ amplitudes for the difficult primary task than for the easy task. The $N_1 - P_2$ amplitudes do not show an interactive effect. Note that the $N_1 - P_2$ amplitudes associated with the easiest primary manual tracking task are smaller than those for the moderate and difficult tasks, except in the case of a-priori condition p=10% on the secondary oddball task.

Figures 4, 5, and 6 show evoked potentials averaged over the 6 subjects for the 3 primary task load difficulties at secondary task probabilities of .10, .20, and .30 respectively. An evoked potential

FIGURE 3. AVERAGE PEAK TO PEAK $N_1 - P_2$ AMPLITUDES OF THE SIX SUBJECTS

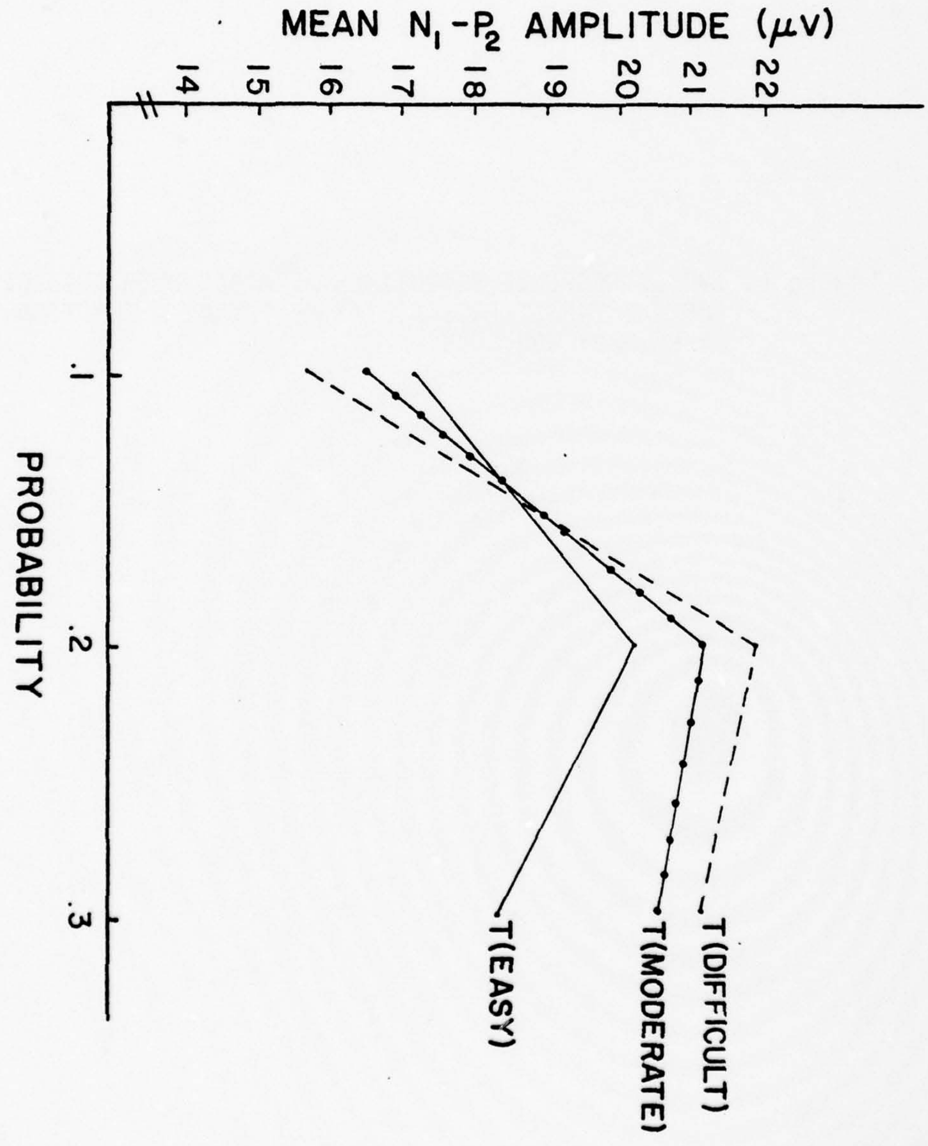

MEAN $N_1 - P_2$ AMPLITUDE (μV)
14
15
16
17
18
19
20
21
22
.1
.2
.3
PROBABILITY
T(DIFFICULT)
T(MODERATE)
T(EASY)

Figure 4. EVOKED RESPONSE POTENTIALS AVERAGED OVER SUBJECTS FOR THE TARGET PROBABILITY OF 10% AS A FUNCTION OF PRIMARY WORKLOAD

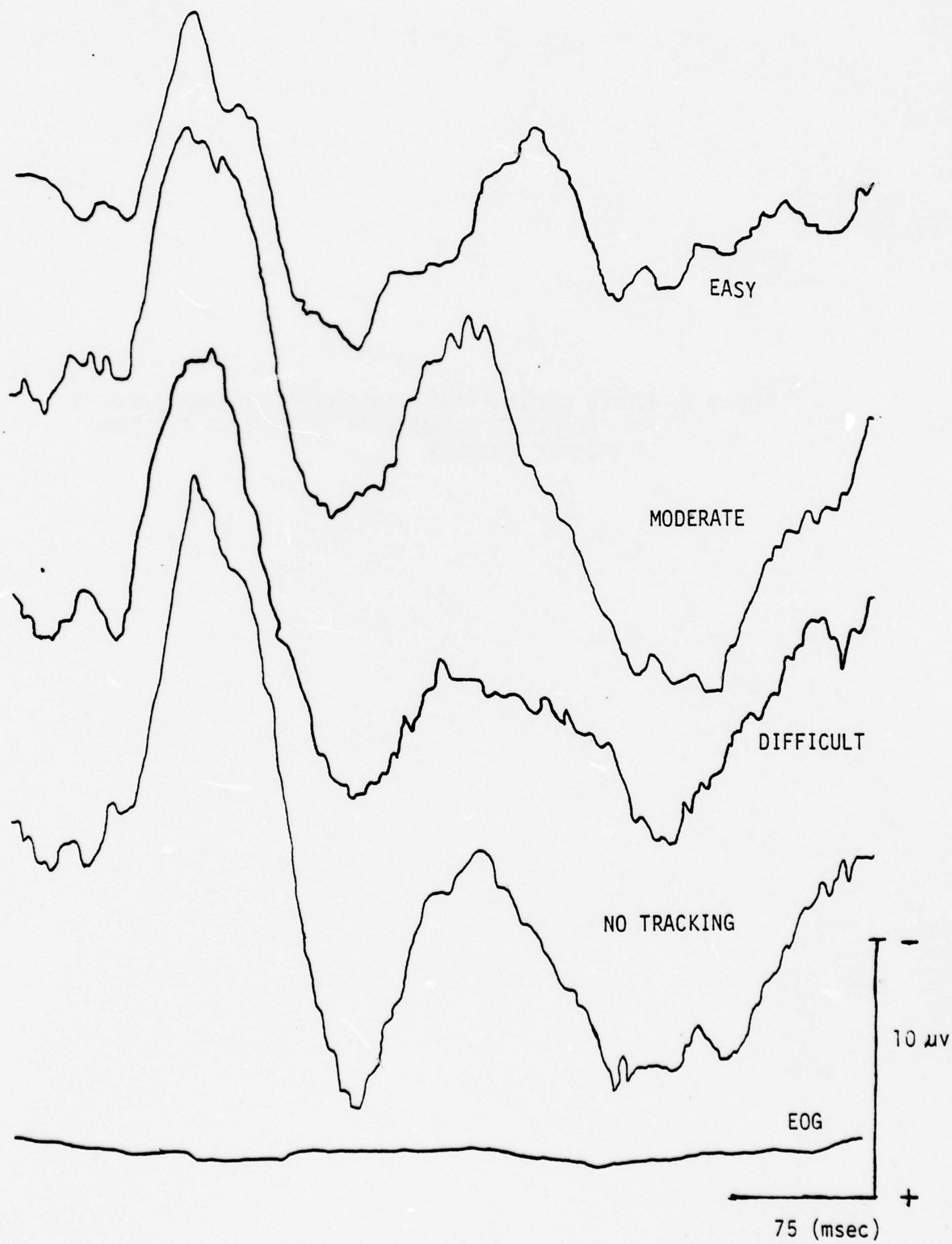

EASY
MODERATE
DIFFICULT
NO TRACKING
EOG
-
10 µv
+
75 (msec)

Figure 5. EVOKED RESPONSE POTENTIALS AVERAGED OVER SUBJECTS FOR THE TARGET PROBABILITY OF 20% AS A FUNCTION OF PRIMARY WORKLOAD

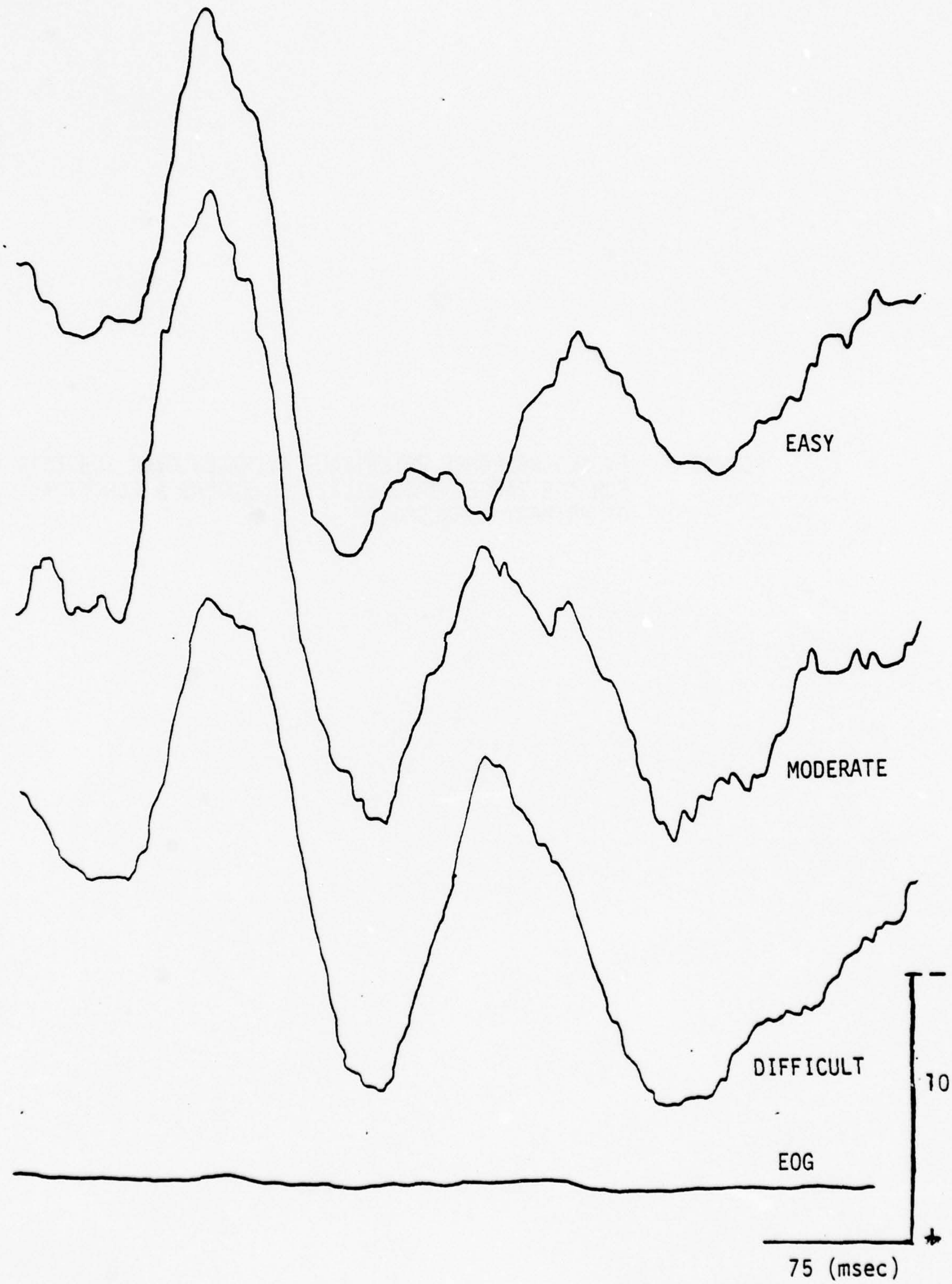
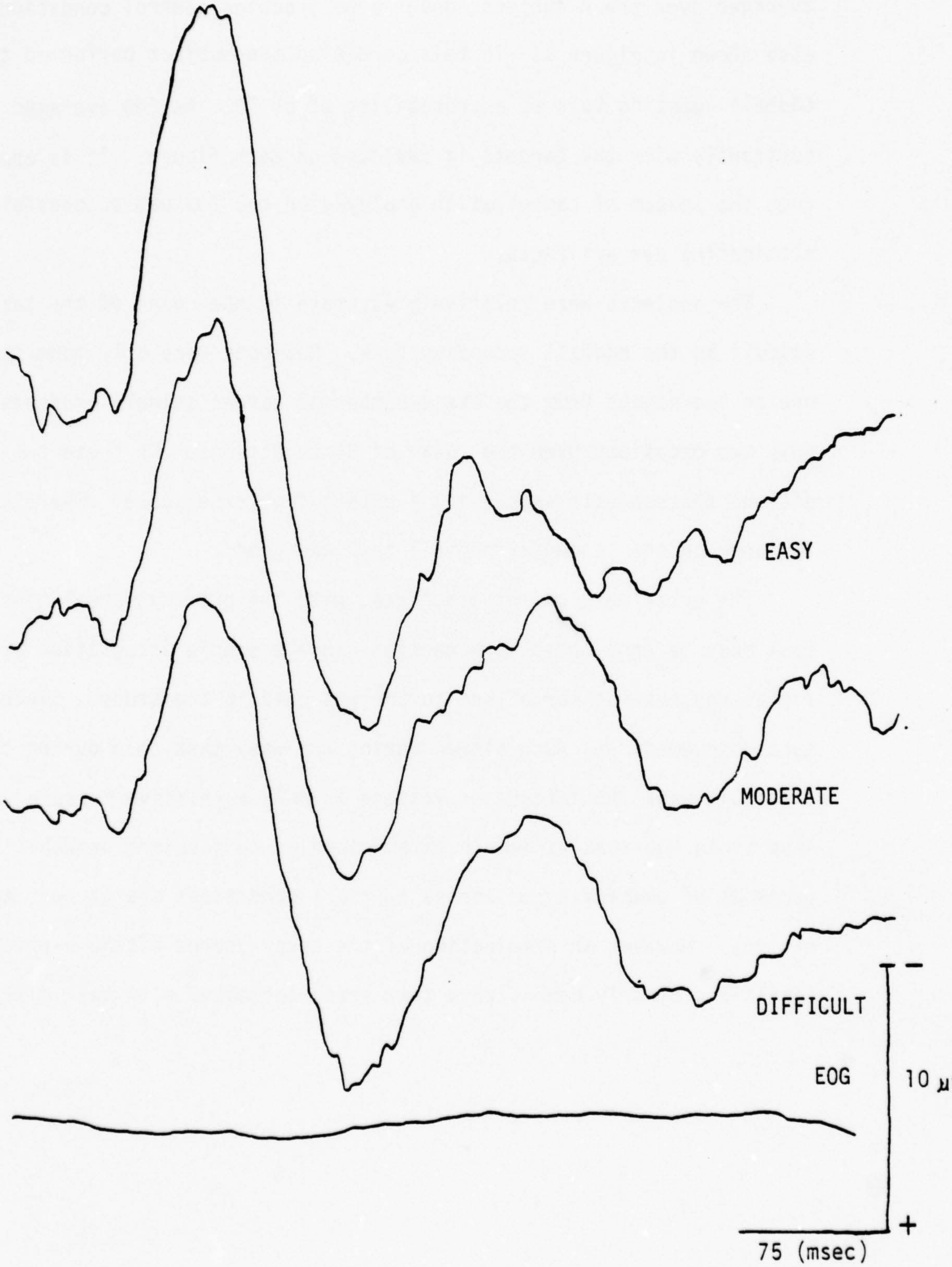

EASY
MODERATE
DIFFICULT
10 µv
EOG
75 (msec)

FIGURE 6. EVOKED RESPONSE POTENTIALS AVERAGED OVER SUBJECTS FOR THE TARGET PROBABILITY OF 30% AS A FUNCTION OF PRIMARY WORKLOAD


EASY
MODERATE
DIFFICULT
EOG
10 µv
75 (msec)

averaged over the 6 subjects under a no tracking control condition is also shown in Figure 4. In this condition the subject performed the oddball counting task at a probability of p=.10. An EOG averaged concomitantly with the targets is depicted on each figure. It is apparent that the system of cancellation employed on the EOG was successful in eliminating eye artifacts.

The subjects were relatively accurate in the count of the target stimuli in the oddball secondary task. Subjects were only more than one or two counts from the exact number of target stimuli presented on only two occasions over the total of 54 conditions. On these two conditions the subjects were 3 and 4 counts from true score. Overall performance on the secondary oddball task was good.

The error measurement associated with the primary manual tracking task must be considered with caution since a single integration of voltage difference between target and cursor was used in the study. Since cursor movement was much slower during the easy task than during the difficult task the integrated voltage is only a relative measure of error. Also since the runs varied in time according to a-priori probability, comparisons of average error across a-priori conditions are at best approximations. However an examination of the error scores within a-priori conditions clearly demonstrate that error increased with task difficulty.

DISCUSSION

The results of the study corroborate the finding of Sutton et al (1965) in that stimulus uncertainty produced a prominent P_3 component. Also the strongest effect in the study was the a-priori probability associated with the secondary task. This result is in agreement with the findings of Donchin et al (1977) and Duncan-Johnson and Donchin (1977). The P_3 component decreases in amplitude with increase in probability of the target stimulus.

Although no test of the effect of random inter-stimulus interval was undertaken; the robust nature of the P_3 component, demonstrated in all experimental conditions, would seem to argue that the temporal uncertainty produced a strong effect on the EEG. McCarthy and Donchin (1976) found a decrease in the P_3 component at the central site (CZ) when temporal uncertainty was introduced in a Bernoulli sequence "oddball" paradigm. However, Ford et al (1976) attributed an increase in P_3 in a sequential task to the temporal uncertainty. McCarthy and Donchin (1976) argue that Ford et al (1976) confounded sequential with global probability since the study was unable to dissociate the multiple effects of stimulus density associated with increase in rate of presentation. Since the results of the present study demonstrate a large P_3 component at the CZ site and that random inter-stimulus intervals were employed, it would seem that Ford et al (1976) were correct in concluding that temporal uncertainty produced differences in the P_3 component. Further work is required to delineate more precisely the effects of temporal uncertainty on event related potentials.

The significant effect of target probability on the N_1 component lends further support to the conclusion of Picton et al (1971) that this particular component although exogenous in nature can be under control of endogenous structures in the cortex. Clearly the N_1 component is not strictly determined by the physical parameters of the stimulus. Hillyard's (1977) proposed model would predict the N_1 changes as a function of target probability if one assumes a change in processing accorded the stimuli under the 3 experimental conditions. Also, if one assumes the hierarchally ordered stages of processing hypothesis defined by Hillyard (1977), the N_1 component should show less change as a function of primary task load than the P_3 component. Although little change was noted on the N_1 component as a function of task difficulty, little change was noted on the P_3 component under similar conditions of the primary task.

It is difficult to explain the demonstrated small changes in the P_3 component as a function of primary task load. The primary tracking task was well practiced so learning effects should have been minimal. Task difficulty was matched to each subject's tracking ability and the error rates as well as subjective reports indicated the tracking levels used were easy, moderately difficult, and extremely difficult. Donchin et al (1977) states that first order changes in P_3 amplitude are not useful indicators of workload and that the amplitude of the P_3 declined precipitously as soon as a tracking task was imposed on the subject. Our results showed no decline in P_3 amplitude even under extreme primary task difficulty. It is possible that the use of both random sequence of tones and random inter-stimulus intervals produced a powerful uncertainty effect that overpowered the primary task effect. An alternative explanation

is that the subject might have adopted strategies that were effective in maintaining a fairly constant performance on the secondary task.

REFERENCES

1. Corby, J.C.and Kopell, B.S., "The Effect of Predictability on Evoked Response Enhancement in Intramodal Selective Attention," Psychophysiology, 1973, 10, 335-346.

2. Courchesne, E., Hillyard, S.A., Hink, R.F., Schwent, V.L., and Picton, T.W., "Stimulus Novelty, Task Relevance and the Visual Evoked Potential in Man," EEG & Clinical Neurophysiology, 1975, 39, 131-143.

3. Donchin, E., Kubovy, M., Kutas, M., Johnson, R., and Herning, R.I., "Graded Changes in Evoked Response (°300) Amplitude as a Function of Cognitive Activity," Perception & Psychophysics, 1973, 14, 319-324.

4. Donchin, E., "The Vocabulary of Brain Potentials: Inferring Cognitive Events from Brain Potentials in Operational Settings," AD-A029452 U.S. Dept of Commerce , August 1976.

5. Donchin, E., Ritter, W., and McCallum, W.C., "Cognitive Psychophysiology: The Endogenous Components of the ERP. Arlie House Conference, 1977.

6. Duncan-Johnson, C.C. and Donchin, E., "On Quantifying Surprise: The Variation of Event-Related Potentials with Subjective Probability," Psychophysiology, 1977, 14, 456-467.

7. Friedman, D., Hakeram, G., Sutton, S., and Fleiss, J.L., "Effect of Stimulus Uncertainty on the Pupillary Dilation Response and the Vertex Evoked Potential," EEG & Clinical Neurophysiology, 1975, 38, 13-19.

8. Hillyard, S.A., "Electrophysiological Assessment of Attentional Processes in Man," Arlie House Conference, 1977.

9. Hillyard, S.A., Hink, R.F., Schwent, V.L., and Picton, T.W., "Electrical Signs of Selective Attention in the Human Brain," Science, 1973, 182, 177-180.

10. Isreal, J.B., Wickens, C.D., and Donchin, E., "P300 Amplitude Changes during a Tracking Task as a Function of Continuous Variations in Tracking Difficulty": Paper Read at Psychophysiology Meetings.

11. Mast, T.E. and Watson, C.S., "Attention and Auditory Evoked Responses to Low-detectability Signals," Perception & Psychophysics. 1968, 4, 237-240.

12. McCarthy, G. and Donchin, E., "The Effects of Temporal and Event Uncertainty in Determining the Waveforms of the Auditory Event Related Potential (ERP)," Psychophysiology, 1976, 13, 581-590.

13. Picton, T.W. and Hillyard, S.A., "Human Auditory Evoked Potentials. II Effects of Attention," EEG & Clinical Neurophysiology, 1974, 36, 191-200.

14. Picton, T.W., Hillyard, S.A., Galambos, R., and Schiff, M., "Human Auditory Attention: A Central or Peripheral Process,"

15. Ritter, W. and Vaughn, H.G., "Average Evoked Responses in Vigilance and Discrimination: A Reassessment," Science, 1969, 164, 326-328.

16. Rohrbaugh, J.W., Donchin, E., and Ericksen, C.W., "Decision Making and the P300 Component of the Cortical Evoked Response," Perception & Psychophysics, 1974, 15, 368-374.

17. Roth, W.T., Ford, J.M., Lewis, S.F., and Kopell, B.S., "Effects of Stimulus Probability and Task Relevance on Event-Related Potentials," Psychophysiology, 1976, 13, 311-317.

18. Ruchkin, D.S., Sutton, S., and Tueting, P., "Emitted and Evoked P300 Potentials and Variations in Stimulus Probability. Psychophysiology, 1975, 12, 591-595.

19. Squires, K.C., Donchin, E., Herning, R.I., and McCarthy, G., "On the Influence of Task Relevance and Stimulus Probability on Event-Related Potential Components. EEG & Clinical Neurophysiology, 1977, 42, 1-14.

20. Squires, N.K., Donchin, E., Squires, K.C., and Grossberg, S., "Bisensory Stimulation: Inferring Decision-Related Processes

from the P300 Component," J. of Exp. Psychology: Human Perception & Performance, 1977, 3, 299-315.

21. Sutton S., Braren, M., Zubin, J., and John, E.R., "Evoked-Potential Correlates of Stimulus Uncertainty," Science, 1965, 150, 1187-1188.

22. Tueting,P., Sutton, S., "The Relationship Between Pre-Stimulus Negative Shifts and Poststimulus Components of the Averaged Evoked Potential," in S Kornblum (Ed.) Attention and Performance IV. New York: Academic Press, 1973.

23. Tueting, P., Sutton, S., and Zubin, J., "Quantitative Evoked Potential Correlates of the Probability of Events. Psychophysiology, 1971, 7, 385-394.

24. Wickens, C.D., Isreal, J., Donchin, E., "The Event-Related Cortical Potential as an Index of Task Work Load," Human Factors Society, San Francisco, .977.

Appendix 1

The initial portion of this experiment is to train you to track a target as accurately as possible. We will start you on an easy level and proceed to a more difficult task level in one minute trials. When you reach a level of difficulty that is extremely hard, we will stop.

We shall then have one more hour of training on the tracking task of different durations with the addition of a secondary task. You will hear a series of randomly occurring hi and lo pitched tones. Your task will be to keep a silent count of the number of occurrences of the hi/lo pitched tone. When the tracking task is completed, you will be asked to report your final count. Try to count as accurately as possible while maintaining your tracking task.

Your task is identical to your last practice session except we will present three levels of tracking difficulty. We will be recording the eye movement and brain potentials from the surface of your head. This procedure will entail cutting a spot of hair on top of your head, rubbing the surface with alcohol, and taping electrodes to the skin.

AN INFORMATION-THEORETIC DESCRIPTION OF THE CxC SYSTEM

Duane G. Leet, Ph.D.
University of Dayton Research Institute
Dayton, Ohio 45469

August, 1978

Aerospace Medical Research Laboratory
Aerospace Medical Division
Air Force Systems Command
Wright-Patterson Air Force Base, Ohio 45433

PREFACE

This report was supported by the Air Force Office of Scientific Research through the USAF-ASEE Summer Faculty Research Program (WPAFB), Contract F44620-76-C-0052, The Ohio State University, Columbus, Ohio.

ACKNOWLEDGMENTS

I would like to thank Dr. J. Ryland Mundie and Capt. Vincent Mortimer for their technical support. Dr. Mundie was the primary designer of the CxC System. Others who contributed to its design and development were Lt. R. Gruenke, Dr. T. Hartrum, Capt. D. Osborne, Capt. R. Rock, Mrs. L. Scofield, and M. Ziskin.

AN INFORMATION-THEORETIC DESCRIPTION OF THE CxC SYSTEM

Duane G. Leet
The University of Dayton Research Institue
Dayton, Ohio

ABSTRACT

The CxC System is a general purpose hybrid hardware and software system for extracting transient patterns from continuous analog, quasi-periodic, audiofrequency signals in real- or near real-time. The design of the system relies heavily on studies of the information processing characteristics of the mammalian auditory system.

This report describes the system from an information-theoretic point of view. A block diagram of the system, structured as the classic preprocessor-feature extractor-classifier decision-theoretic pattern classifier, is presented and each of the component information transformations described. The preprocessor contains a preamplifier, a Middle Ear circuit, and an Analog Cochlea. The latter is a nonuniform, reflectionless, leaky transmission line with a number of useful information processing characteristics. The feature extractor contains two levels of neuromime networks that together provide instantaneous component period and phase information about the input signal. The classifier contains hardware that compares ten millisecond intervals of the feature extractor response in real-time with up to sixteen stored reference patterns. It also contains software to decide when a decision can be made and what the input pattern sequence is.

TABLE OF CONTENTS

LIST OF FIGURES

SECTION 1
INTRODUCTION

The classic implementation of a decision-theoretic pattern recognition system consists of a Preprocessor to condition the raw signals, a Feature Extractor to extract relevant information and to generally reduce the dimensionality of the problem, and a Classifier to provide either a definitive classification, an ordered list of classes- the order based on Classifier-computed probabilities - or the probabilities themselves (Figure 1). A principal design objective for the Feature Extractor is to reduce the incoming information to exactly that which is essential for recognition- no more, no less. This is done through one or more levels of hardware or software components that successively extract features from the signal. From the digital processing point of view, each successive level reduces the number of bits required to represent the input signal, with each successive bit becoming more important. It isn't too often that this objective is completely satisfied. Indeed, to many people the problem being solved is considered trivial or uninteresting if this objective is met.

A common problem is that the relevant pattern features are neither obvious nor known. There are two practical consequences of this. First, the design approach tends to be conservative, with more information passed to each level than is necessary to make the decisions required at that level. The Estimator, in turn, must be made overly complex in order to separate the potentially significant data from the insignificant data. As a result, not only are the initial system development costs inflated, perhaps leading to a decision to cancel the development of what would have evolved into a useful, cost-effective system, but the system might pass through a series of costly modifications before achieving its design objectives. Second, there is no assurance that the Classifier will perform correctly. The design can be verified only through imperical testing.

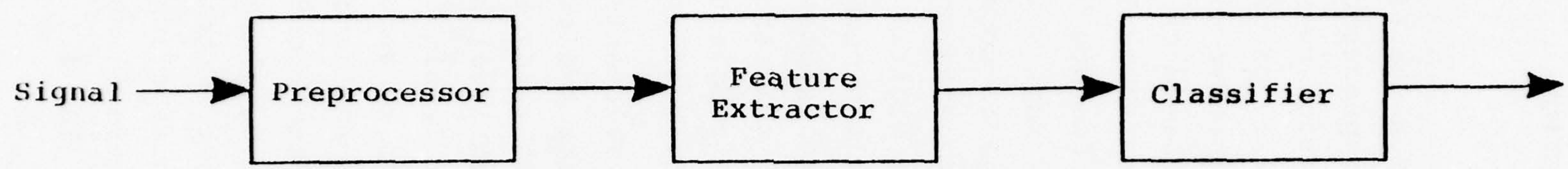


Figure 1. The Decision-Theoretic Pattern Recognition System.

The recognition of isolated and continuous speech are two interrelated but separable problems that are at the most difficult and yet most interesting end of the spectrum of pattern recognition problems. The operational objective in both cases is to be able to produce a transcription of a spoken message without human intervention. In the case of isolated speech the message is a word or short phrase; in the case of continuous speech the message length is indeterminant and, no matter how the message is partitioned, the segment waveforms are generally context sensitive.

Appendices A and B briefly describe the state-of-the-art in these two technologies. The important characteristic of both designs is the comparison of feature pattern vectors with reference pattern vectors, with the features being selected by educated guess and then determined acceptable by empirical test.

The CxC (say "C-squared") system is a general purpose hybrid hardware and software system for extracting transient sequential patterns from continuous analog, audiofrequency signals in real- or near real-time (Speech is an example of such signals.). The design of the system relies heavily on studies of the information processing characteristics of the mammalian peripheral auditory system. Figure 2 is a block diagram of the system, emphasizing its information-theoretic organization. This paper describes the information transformations that are performed by each of the components.

The primary obstacle to acceptance of the CxC system is the engineering community is that the feature vectors generated by the Feature Extractor corresponds to features that are new to it. Furthermore, an explicit mathematical/functional description of the total Feature Extractor transformation does not exist. Consequently, a major portion of this paper is devoted to providing an intuitive understanding of the Feature Extractor's operation and unique capabilities.

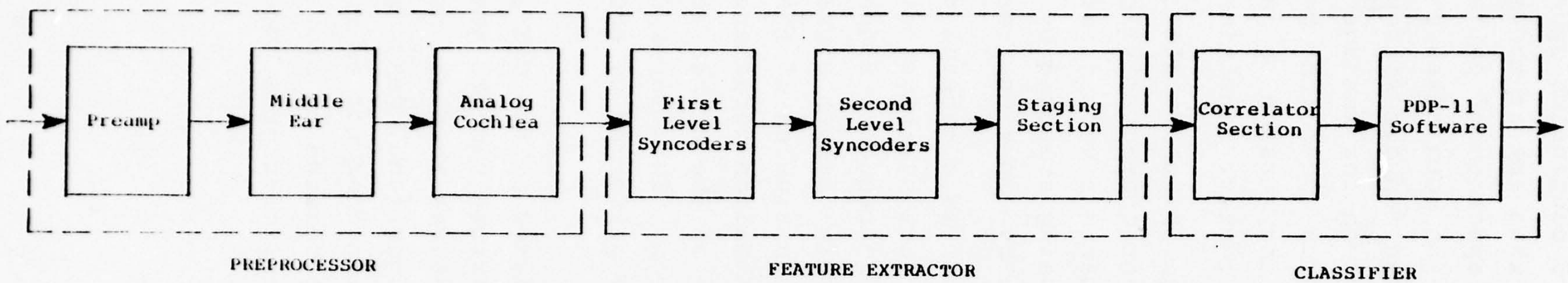


Figure 2. A CxC System Block Diagram Emphasizing Its Information Processing Characteristics.

SECTION 2

INFORMATION PROCESSING CHARACTERISTICS OF THE PREPROCESSOR

Three CxC system components are identified as Preprocessor components: a Preamplifier circuit, a Middle Ear circuit, and the Analog Cochlea. The first two circuits together form a bandpass filter whose transfer function is centered at about 2 kHz, with a relatively wide bandwidth of about 3.5 kHz, and a 6 dB/octave low frequency skirt and a 12 dB/octave high frequency skirt. These components were specifically designed for speech processing, and can be by-passed when signals from other sources are to be processed.

The Analog Cochlea is a nonuniform, leaky, reflectionless transmission line tapped at 48 locations; each tap is called an output channel of the cochlea. Abstractly, this device transforms a scalar function of time into a vector function of time: (amplitude) x (time) ---→ (distance along cochlea) x (amplitude) x (time); it transforms two dimensions into three.

It is important to understand that the Cochlea was not designed to preserve a waveform, as is the usual objective of transmission line design. Instead, it was designed to deliberately alter the signal, and it is the alterations that are important.

The transfer function between Preamp input and an output channel is that of an asymmetric bandpass filter with a steep 100 dB/octave high frequency skirt and a 6 dB/octave low frequency skirt. As demonstrated in Figure 3, each output channel has a different center frequency, with the higher center frequency channels located closer to the Cochlea's input.

A signal requires a finite amount of time to travel down the Cochlea, with the speed of a wavefront being inversely proportional to the logorithm of its distance (not frequency!) down the Cochlea (Figure 4).

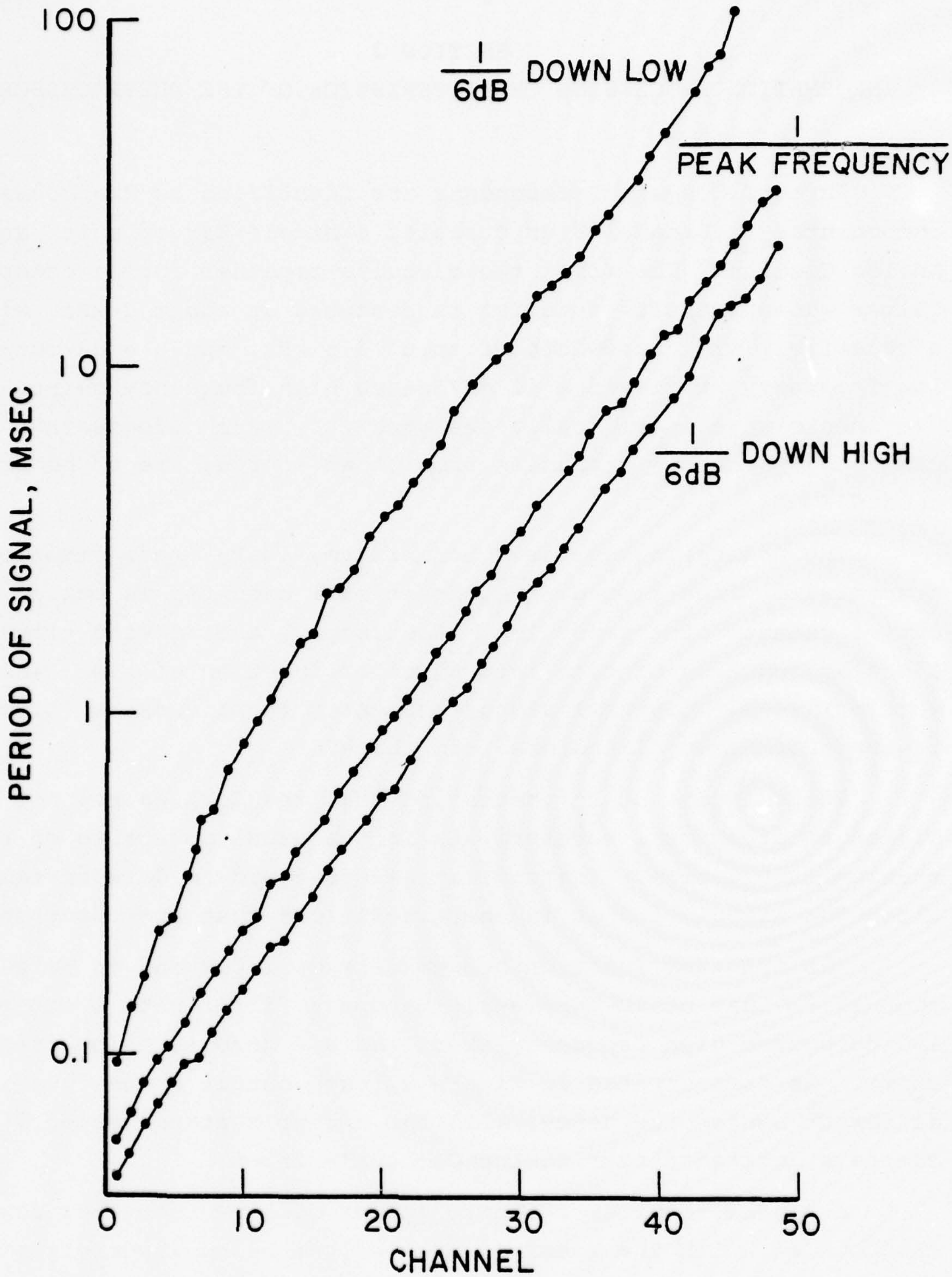


Figure 3. The Cochlear Filter Pass-Bands.

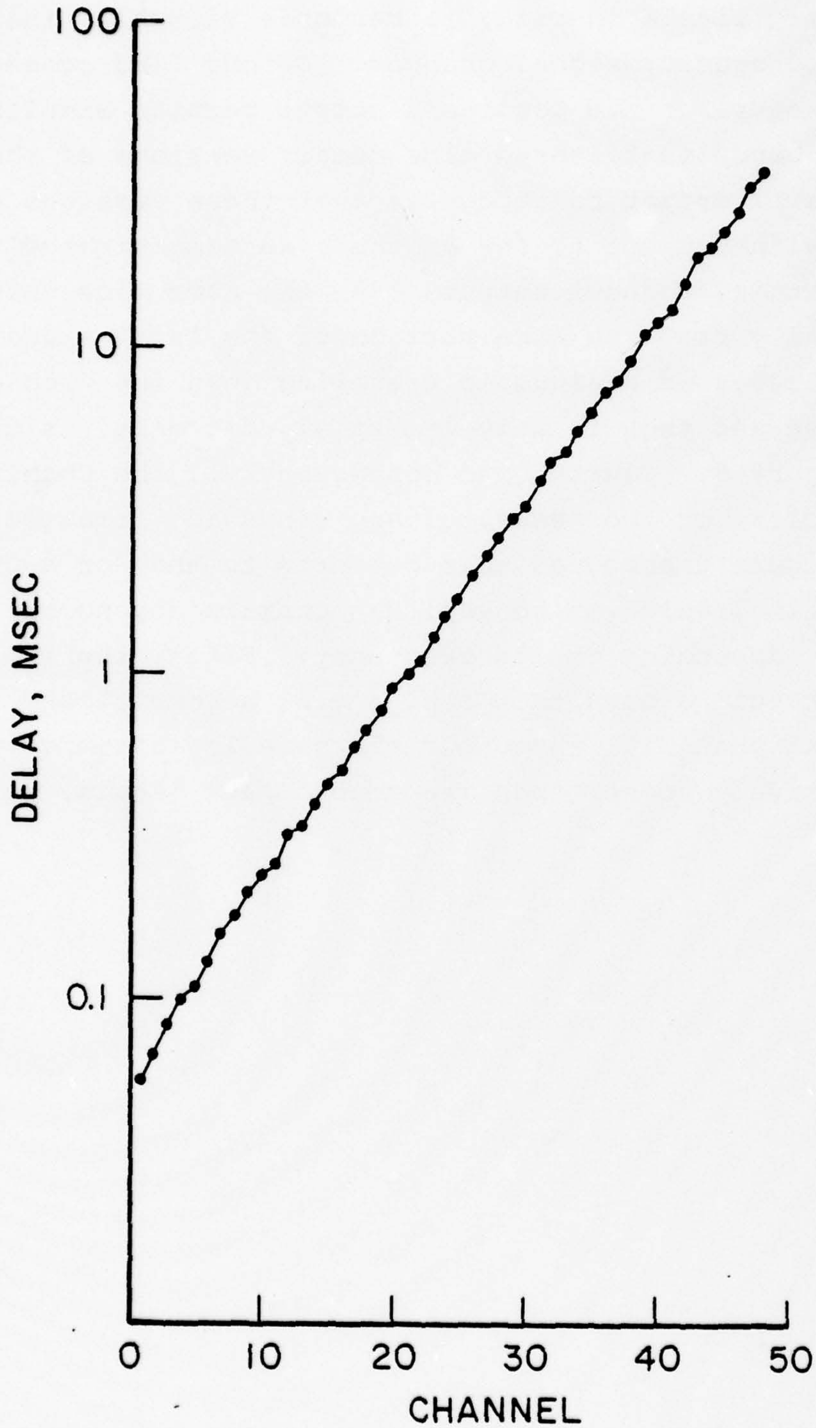


Figure 4. The Cochlear Filter Signal Delay from Input to Channel Output.

This design has a number of useful consequences. First, the Cochlea has very little inertia; it responds virtually instantaneously to audiofrequency signal changes. Second (and consequently), the parallel nature of the Cochlea's output permits simultaneous comparison of bandpass-filtered time domain versions of the input signal. It is important to recognize that these versions are also time-shifted with respect to one another, as demonstrated in Figure 5. That is, channel outputs over the same time intervals are not generally from the same portion of the input signal. Third, the peak amplitudes of a sinusoid traveling down the Cochlea will slowly increase and then relatively rapidly decrease, as demonstrated in Figure 6. Fourth (and consequently), the Cochlea stores about 1.75 cycles, or two peaks, of any sinusoid, irrespective of frequency. Figure 6 compares this property to that of a uniform transmission line, which in general can contain any number of peaks of a sinusoid, depending on its frequency. Fifth, the two peaks of sinusoid traveling down the Cochlea will become closer together, since the first peak will encounter the same logarithmic velocity gradient relatively sooner than the second peak (Again, refer to Figure 6.).

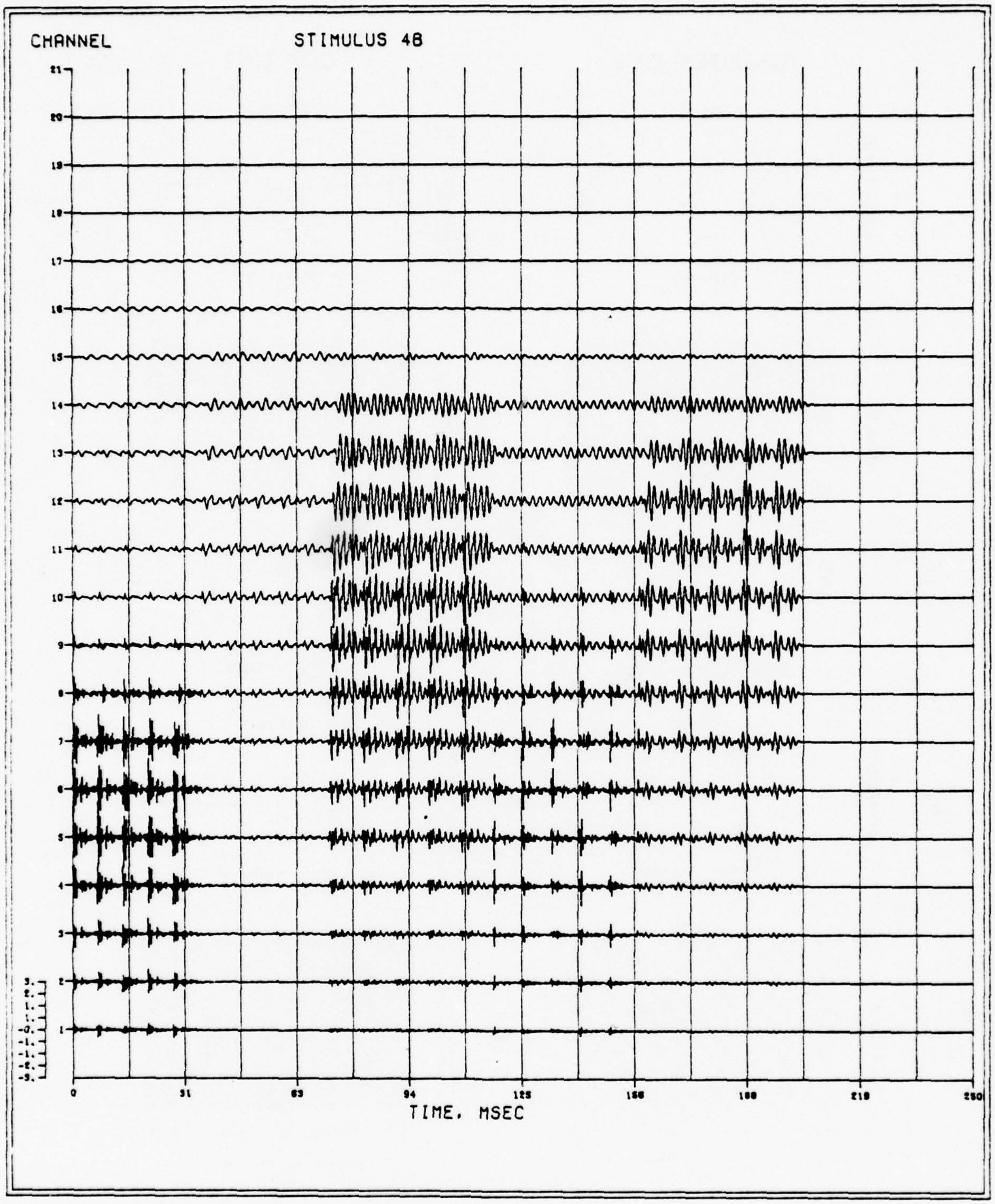


Figure 5. The Cochlear Surface Generated by the Cochlear Filter in Response to a Sequence of Five Vowels.

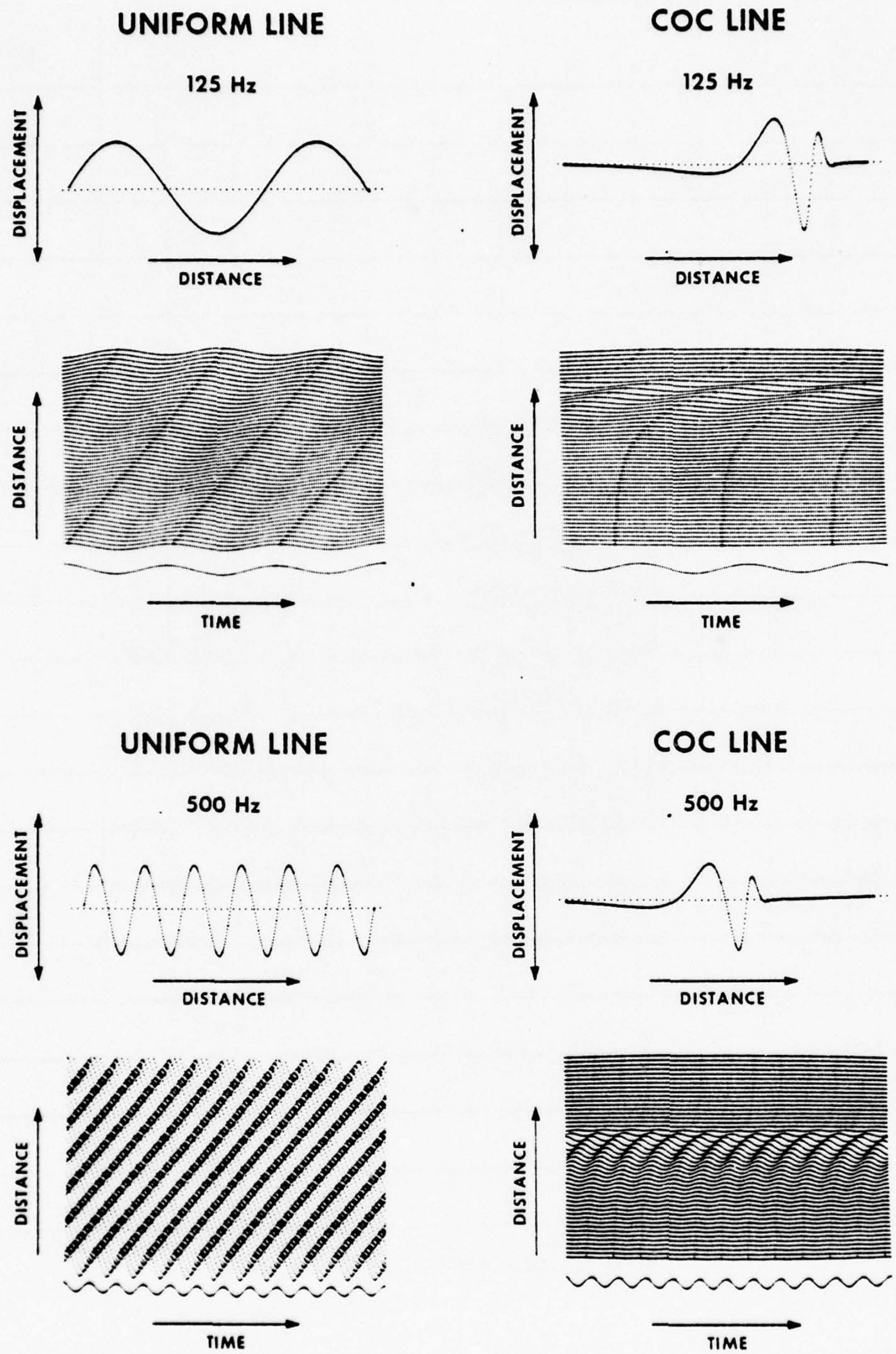


Figure 6. Comparison of Signal Processing Properties of a Conventional Uniform Transmission Line with a Transmission Line Which Models the Inner Ear.

SECTION 3

INFORMATION PROCESSING CHARACTERISTICS OF THE FEATURE EXTRACTOR

The Feature Extractor has three information processing components, the Primary Level Syncoder Network, the Secondary Level Syncoder Network, and the Staging Section, all implemented in hardware. The overall information transformation performed by the Primary and Secondary Level Networks will be described first. Examples of their behavior are provided. The operation of the Staging Section will then be discussed.

At the abstract algebraic level, the two Syncoder Networks and their supporting hardware and software perform an overall transformation from a 48-dimension real vector space to a 32-dimension binary vector space, with the binary vectors generated at 5 microsecond intervals. The transfer function is output signal dependent. It has not been described analytically, although it has of course been implemented in hardware and can be simulated in software.

The overall transformation for 30 of the 32 channels* can be described operationally in the following way: Each channel is sensitive to a range of <u>periods</u>. If at any instant the Cochlea stimulus contains detectable frequency components** for which 1/frequency is in that range of periods, then one pulse per period will be generated, the location of the pulse within the period being a function of the magnitude of the components in that range.

---

*The other two channels have special objectives. One was designed specifically for speech recognition: it is a pitch period marker. The other generates a pulse train whose density is logarithmically proportional to the Cochlea input signal's magnitude.

**These are instantaneous frequency components that are the inverse of the time between two identical points on a waveform.

One of the design objectives was to have each channel be sensitive to a different range of periods, with a slight overlap in range between adjacent channels. Figure 7 demonstrates the actual period ranges, in terms of the frequencies of a sinusoidal stimulus. As can be seen, the objective was, for the most part, met.* Figure 8 demonstrates this property in a different way. A pair of Wavetek VCGs were arranged to generate a stimulus that maintained a constant frequency of 125 Hz for 20 ms and then linearly increased in frequency to about 500 Hz during the next 20 ms. Figure 8 is the response of the Secondary Level to this stimulus.

Channel period sensitivity in CxC is a function of stimulus magnitude. Figure 9 demonstrates this property. The Wavetek VCGs were arranged so that a 550 Hz sinusoid was modulated by a 25 Hz sinusoid. The maximum peak-to-peak magnitude of this stimulus was about 8.2 volts; the minimum was about 0.6 volts. Figure 10 shows that from zero to five channesl will produce pulses over this range of stimulus magnitudes. The range of magnitudes for which exactly one channel generates a pulse with this stimulus was found in a separate experiment to be from 1.5 volts to 2.0 volts peak-to-peak.

As a result of these characteristics, the Secondary Level's overall response pattern is sensitive to the "instantaneous" time domain waveform. (Parenthetically, the frequency domain interpretation of this property is that the Secondary Level's overall output contains both "instantaneous" phase and frequency information. Although this frequency domain terminology has some meaning in the abstract sense, practically it is generally not possible to generate instantaneous frequency and phase information from a signal. That is the reason frequency domain terminology is being avoided in this paper.) To demonstrate this, one Wavetek VCG

*That this objective was not entirely met was due in part to a second objective: The components in the Feature Extractor were to model the information processing characteristics found in the mammalian auditory system. The syncoders and their associated networks represent a new technology with operating characteristics that are not yet fully mastered.

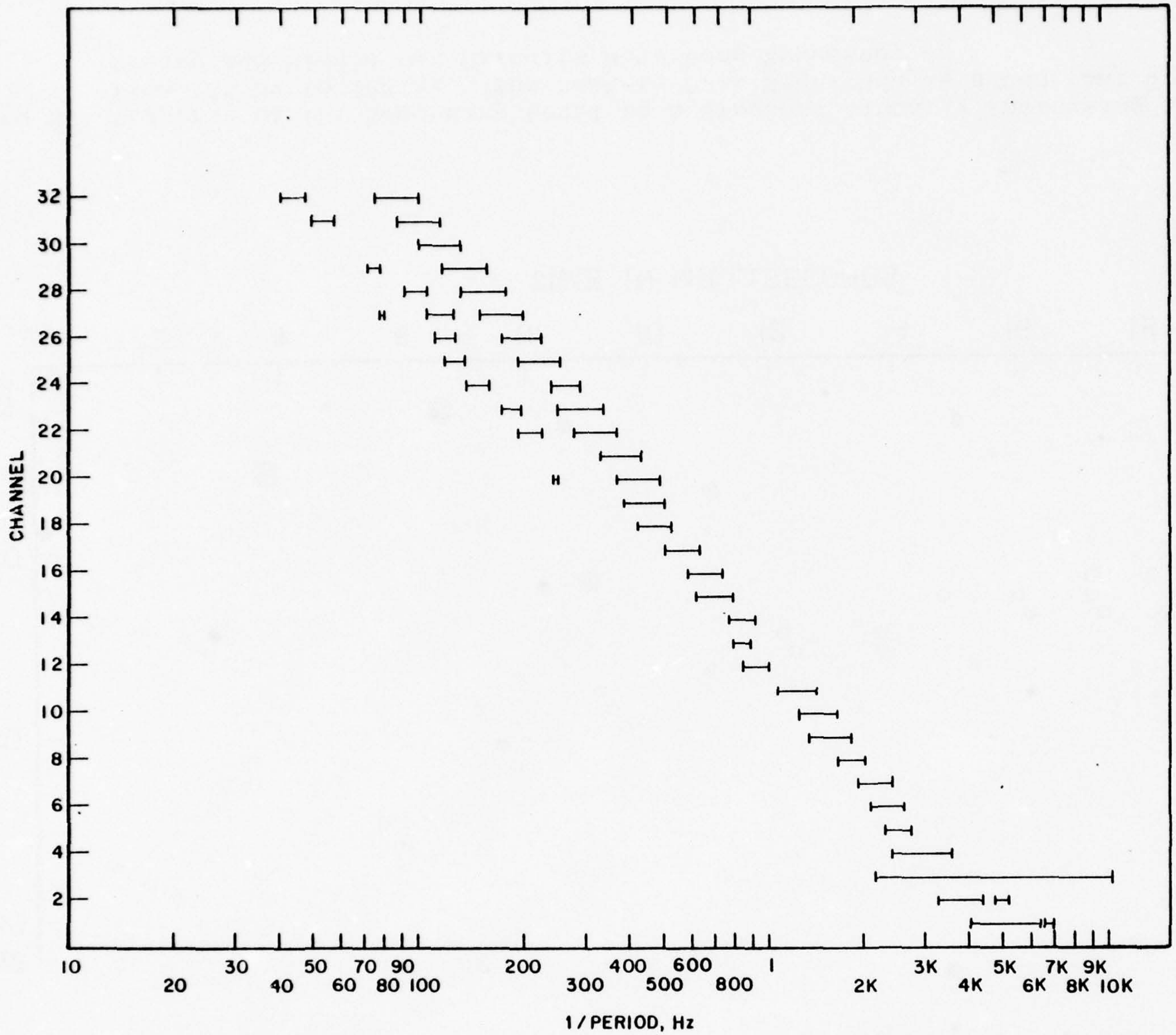


Figure 7. CxC Secondary Level Period Sensitivity With Middle Ear Bypass. The Amplitude at the Cochlear Filter Input Is 1.8 V P-P.

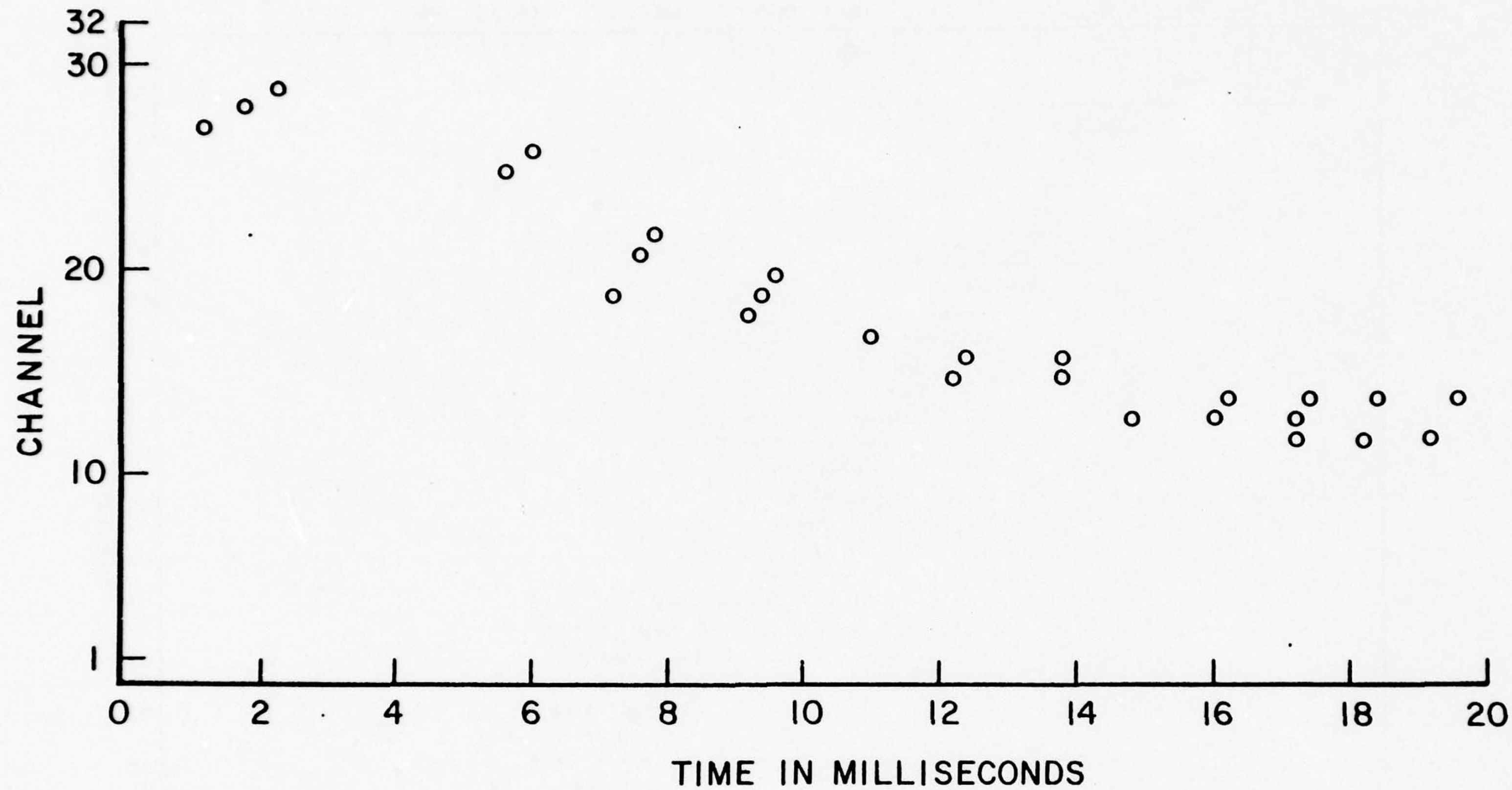


Figure 8. Response of the Secondary Level to a Sinusoid Linearly Increasing in Frequency from 125 Hz to 500 Hz. The Peak-to-Peak Magnitude is a Constant 0.6 V. The Preamp and Middle Ear Circuits Have Been Bypassed.

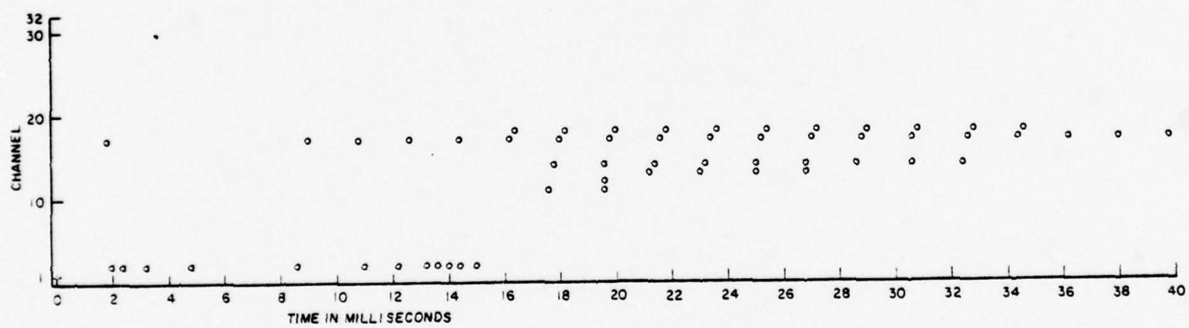


Figure 9. Secondary Level Response to an Amplitude Modulated 550 Hz Sinusoid. The Modulating Signal is a 25 Hz Sinusoid. The Maximum Peak-to-Peak Amplitude of the Stimulus is About 8.2 Volts; the Minimum Peak-to-Peak Amplitude is About 0.6 Volts. The Preamp and Middle Ear Circuits Have Been Bypassed.

generated a 333 Hz sinusoid and the other generated a 499 Hz sinusoid. One of the Waveteks had a phase-locking feature, permitting the phase relationship of the common fundamental, 166 Hz, to be adjusted. Figure 11 show three different stimuli constructed from these two sinusoids, along with the Primary and Secondary Level responses. It can be seen that the pulse patterns of both the Primary and Secondary Levels are unique for each of the three stimuli.

The Staging Section of the Feature Extractor can be thought of as a bank of 20 shift registers.* Each 5 microseconds the current Secondary Level response vector is shifted into one end of the shift registers.** Since each shift register is 2000 bits long, the register bank can hold the most recent 10 ms of Secondary Level response.

*In the current version, only channels 1-20 are passed on to the Staging Section. The descriptions of the Staging Section and the Correlator Section provided here do not describe how the operations are actually performed, but what the net effect is.

**Optionally, the Primary Level can be input to the Staging Section.

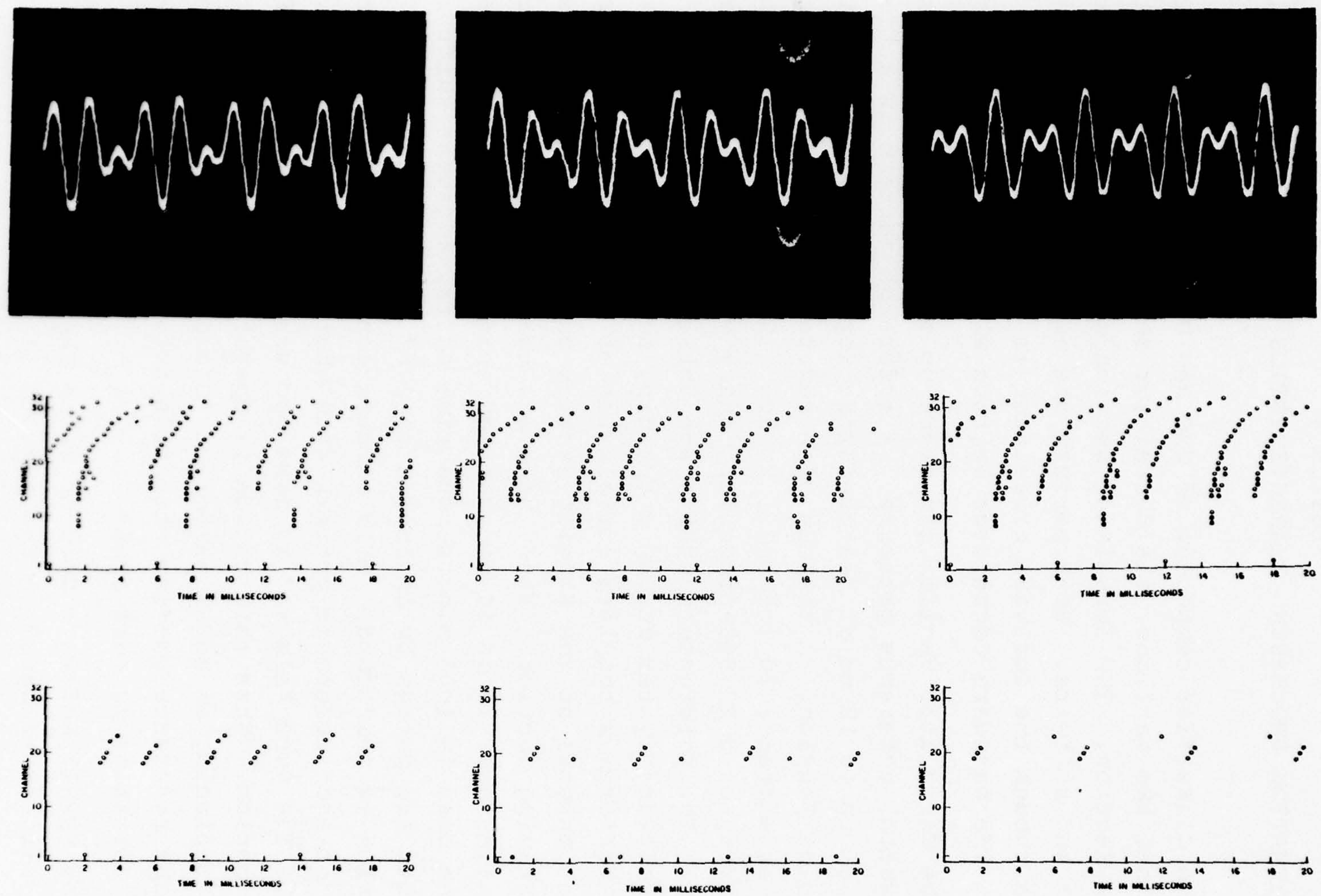


Figure 10. Three Stimuli Constructed From Two Sine Waves of Equal Amplitude (First Row of Diagrams) With a Frequency of 333 Hz and the Other With a Frequency of 499 Hz, and the Associated CxC Responses at the First (Second Row of Diagrams) and Second (Third Row of Diagrams) Levels.

SECTION 4
INFORMATION PROCESSING CHARACTERISTICS OF THE CLASSIFIER

The Classifier component of CxC has been conceptually partitioned into two sections, the Correlator Section and the PDP-11 Software Section. The Correlator Section is implemented in both hardware and software. Both perform the same effective computations, although the software correlator is thousands of times slower. The hardware Correlator Section will be emphasized here.

The Correlator Section can store up to 16 (20 in the software version) reference patterns. A reference pattern is a time segment of up to 10 ms of Feature Extractor response as found in the Staging Section. A maximum size reference pattern can be considered as either a 20 (channel) x 1000 (10 microsecond time unit) binary array or a 20,000-dimension binary vector for classification purposes. The reference patterns are obtained using a special PDP-11 software system during a training session. The Correlator Section performs a complete correlation of the reference patterns with the contents of the Staging Section once every shift period of the Staging Section. Note, however, that since the reference pattern time base is in 10 microsecond units and the Staging Section time base is in 5 microsecond time units, adjacent shift register bits are paired by an inclusive "OR" function before the correlation is performed. This correlation is almost the inner product of each reference pattern with whatever is in the Staging Section. The word "almost" is used because of the existence of a "smear" option. When this option is operating, the effect is the same as replacing an equal number of "0"s on either side of every "1" in the reference pattern with "1"s before the standard inner product operation is performed.

The PDP-11 Software Section makes the decisions required of the overall system based on the results of the correlation. The software in this section must be tailored to the signal transmitter

and pattern generator characteristics. A speech recognition system organization for the PDP-11 Software Section is described in Leet (1978).

REFERENCES

Klatt, Dennis H. (1977), "Review of the ARPA Speech Understanding Project," J. Acoust. Soc. Amer. 62(6):1345.

Leet, Duane G. (1978), "Preliminary Evaluation of the CxC System As an Automatic Speech Recognizer," UDR-TR-78-79, University of Dayton Research Institute, Dayton, Ohio.

Martin, T.B. and E.F. Grunza (1974), "Voice Control Demonstration System," AFAL-TR-74-174, Air Force Avionics Laboratory, Wright-Patterson Air Force Base, Ohio.

APPENDIX A
AN ISOLATED WORD RECOGNITION SYSTEM

Threshold Technology, Inc. is marketing an isolated word recognition system that is proving successful in industrial applications. A version of the system has also been tested in a laboratory cockpit-like environment, with the subjects using a standard microphone and breathing compressed oxygen through a standard oxygen mask (Martin and Grunza, 1974). The overall recognition accuracy was 97.15 percent in this situation. It is instructive to examine the information-theoretic design of this system. The Preprocessor in Figure A-1 extracts 32 features from the input signal. These features, including short term spectral components and zero-crossing rates, were selected through extensive trial and error experimentation. What happens from this point on can best be described by quoting from Martin and Grunza's report:

> The 32 feature lines of the preprocessor are sampled once every 2.5 ms for the duration of the utterance. Detection of the end of the utterance initiates a time normalization routine which divides the data stored in the buffer into 16 equal time segments. The features present within each time segment are examined to insure that they meet minimum duration requirements. Those features which are present for at least this minimum period are assigned a value of 1 for the time segment; those which do not meet this requirement are given a value of 0. Thus, the features present during each time segment may be described by 32 data bits. The complete utterance is then characterized by a 16 x 32 bit array. Interrupts are enabled immediately after the formation of this array so that the data buffer is again ready to accept verbal input.
>
> Ten arrays are stored for each of the vocabulary words during the training phase. The ten arrays are reduced to a common reference array which is then stored in the reference array library. A template threshold factor of 40 percent is used during the generation of the reference array. Thus, a feature must be present in a given time segment for at least 40 percent of the training samples for that feature to have a value of 1 in the reference array. Those features which do not meet this requirement are given a value of 0 in this array.

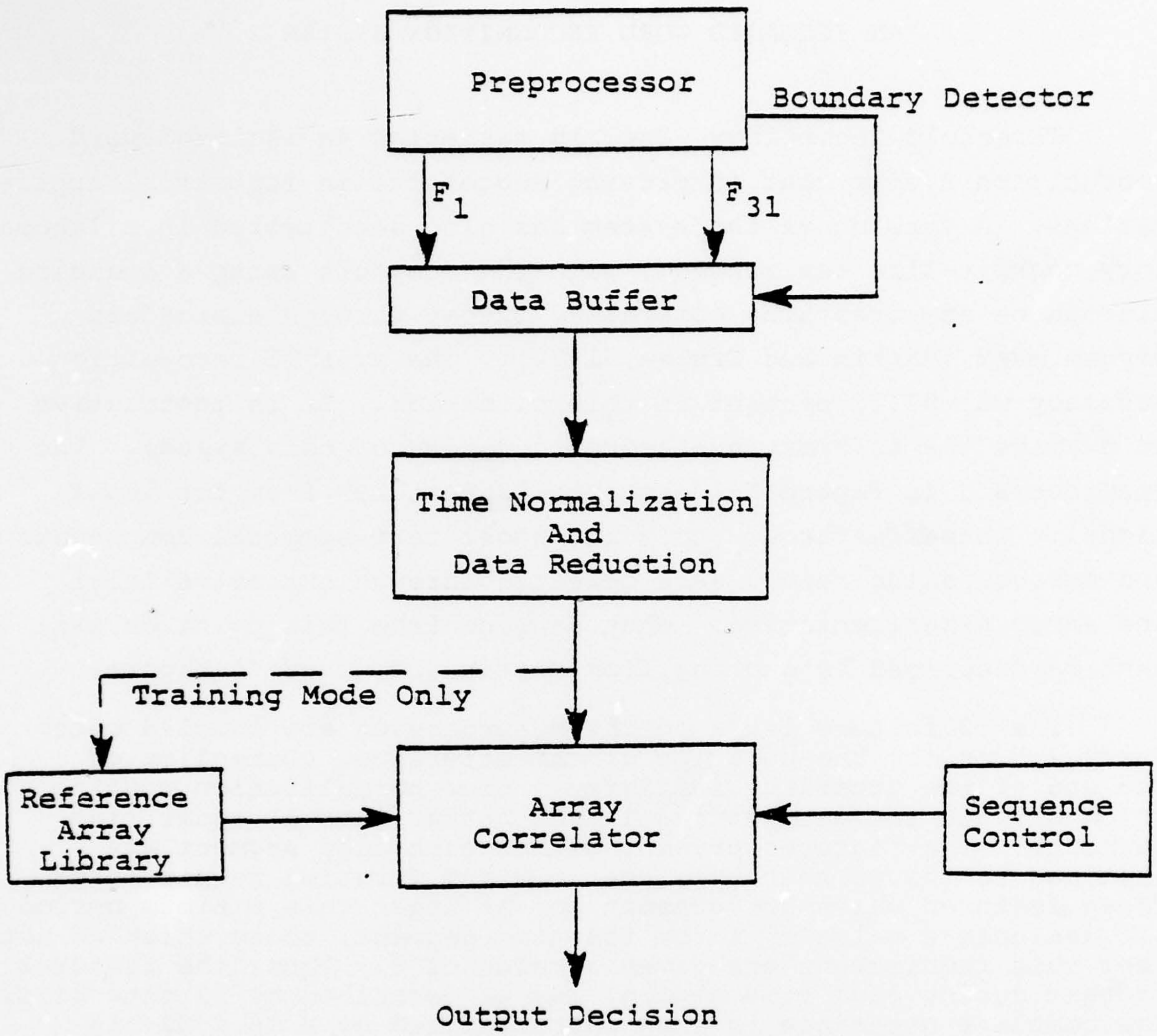


Figure A-1. Block Diagram for Word Recognition Processing Within the Minicomputer of the Threshold Technology, Inc. Voice Control System [From Martin and Grunza (1974)].

In the operating mode, the array resulting from each input utterance is compared against the reference arrays for all currently valid vocabulary words. The vocabulary control data is supplied by the sequence control routine which will be described later. Three separate correlation products are calculated for each valid reference array. These correlations represent the directly superimposed and the shifted plus and minus one time segment conditions. The results of three correlations are combined to obtain a single correlation value for each valid vocabulary word. The word producing the highest correlation score is selected as the input utterance if this correlation score exceeds a present threshold. A reject indication is output if none of the correlation scores exceed the threshold.

APPENDIX B
A CONTINUOUS SPEECH RECOGNITION SYSTEM

Attempts to solve the continuous speech recognition (CSR) problem have been less successful. Figure B-1 is a system diagram of a CSR system, based on a CSR system proposed by Klatt (Klatt, 1977), who, in turn, based his ideas on the most successful of the experimental CSR systems, Carnegie-Mellon University's Harpy system. The speech input in the Figure is a transformed version of the signal generated by the speaker, modified by both noise and linear (hopefully) filters. It is processed by the Front End component, which partitions the signal into segments and generates a multi-dimensional parameter vector for each segment. A priori information about the sentences that can be spoken and how they might be spoken is contained in the Acoustic Segment Lexical Decoding (ACLD) Network. In a separate off-line one-shot generation, the speaker was asked to speak about twenty sentences carefully chosen to contain all the different segment acoustic forms that could occur during the speaker's conversation. Parameter vectors are formed from these segments and placed in a special reference file, the Diphone Dictionary. Each node in the ACLD Network points to a parameter vector in the reference file.

The Search Strategy component within the Bottom End section parses the acoustic segment sequence using the ACLD Network. The a priori probabilities dictate the order in which node exist paths are tested. Each destination node reference pattern is compared to the next acoustic pattern vector using a distance measure. A decision function combines the a priori and a posteriori information to select the best path through the Network. When an end-of-sentence mark is detected in the Network, the candidate sentence is sent to the Search Strategy component in the Top End section. It uses an Augmented Transition Network with Semantics to decide whether the sentence is in acceptable form. If it is, the sentence is output; otherwise, the Bottom End Search Strategy component is notified. It locates the next best path through the Network,

starting from the beginning of the acoustic segment sequence, and this is evaluated by the Top End section. This process is repeated until an acceptable sentence is identified or the Network possibilities are exhausted.

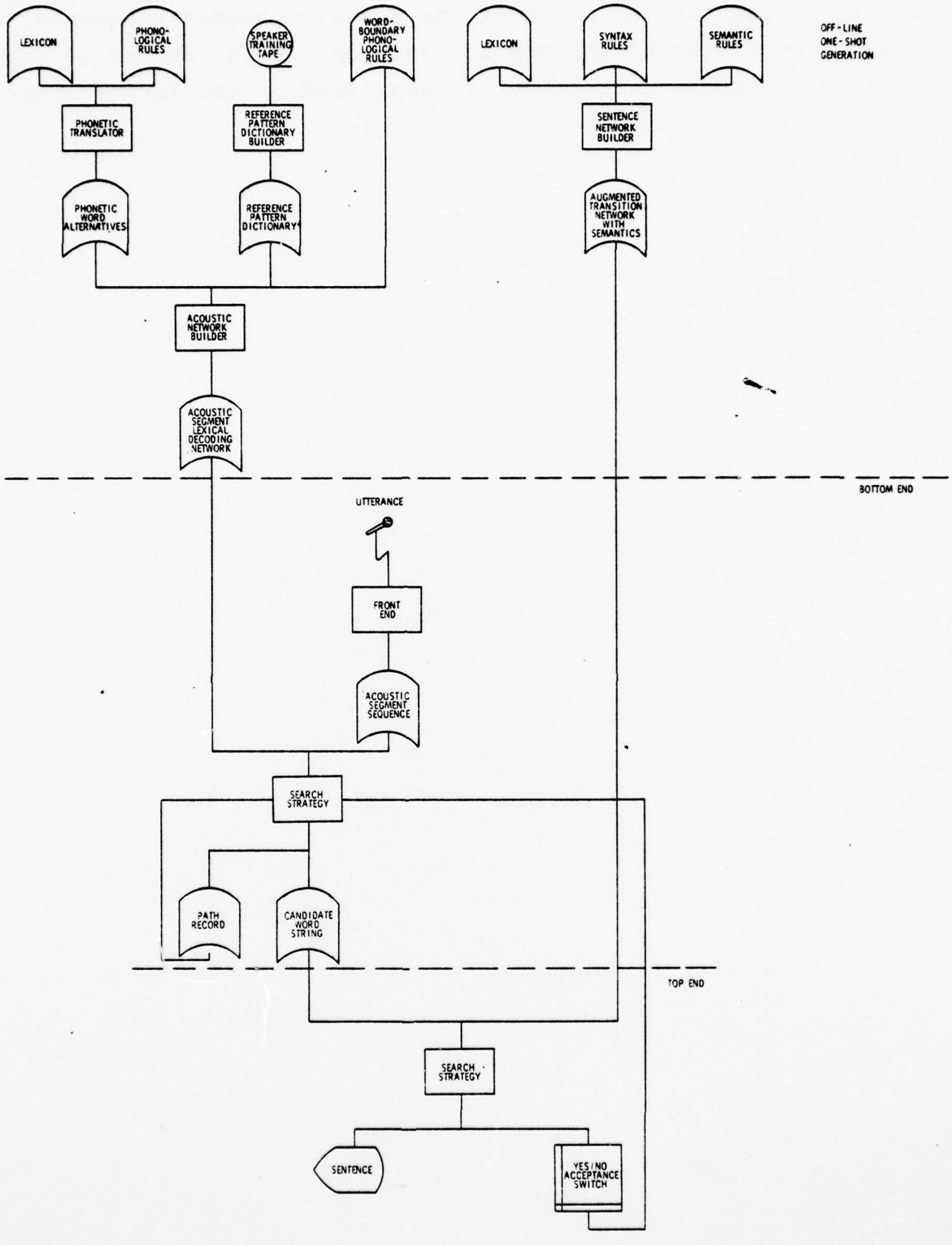


Figure B-1. A System Diagram of a CSR System, Based on a CSR System Proposed by Klatt (1977).

THE EFFECTS OF BROMCHLORODIFLUOROMETHANE (1211) ON CANINE PURKINJE FIBERS

S. Mark Strauch and Capt. Richard A. Davis

AMRL/THT
Wright-Patterson AFB

THE EFFECTS OF BROMCHLORODIFLUOROMETHANE (1211) ON CANINE PURKINJE FIBERS

S. Mark Strauch*

The Ohio State University, Columbus, Ohio

Abstract

Bromchlorodifluoromethane (Compound 1211) is used by the United States Air Force as a fire extinguishing agent. Halogenated alkanes, including 1211 and other compounds, are known to have certain cardiotoxic actions, including sensitization of the heart to catecholamines. 1211 was tested for its effects on the membrane electrical characteristics of isolated perfused canine Purkinje fibers. This compound produces changes in the Purkinje action potential (AP) at concentrations attainable by breathing less than 5% (v/v) in air. The most pronounced effect was on the AP duration and effective refractory period. Additional studies with 1211 and isoproterenol showed that this compound protentiated the effects of isoproterenol, providing evidence for a direct sensitizing action on the heart. The results of these experiments provide a basis for discussing the arrhythmogenic action of 1211.

* Instructor, The Ohio State University, College of Veterinary Medicine, Department of Veterinary Physiology and Pharmacology, Columbus, Ohio.

INTRODUCTION

Cardiac sensitization was first demonstrated by Levy (1911) over sixty years ago when he showed that an injection of epinephrine caused ventricular fibrillation in cats subjected to chloroform inhalation. Subsequent studies with other halogenated alkanes, including Freons and inhalation anesthetics, have shown this class of compounds to be potent sensitizers to the arrhythmogenic action of catecholamines (Aviado, 1974; Back and VanStee, 1977). That serious cardiac arrhythmias may result from the inhalation of these compounds is evidenced by the epidemic of "sudden sniffing deaths" that occurred in this country in the late 1960's among abusers of Freon aerosol propellants (Bass, 1970; Harris, 1973). Several low molecular weight halogenated alkanes are of special interest to the United States Air Force because of their use as fire extinguishing agents (VanStee, 1974). Of the compounds currently in use, compound 1211 ($CBrClF_3$, bromochlorodifluormethane, BCP) was selected for testing in the series of experiments presented here. 1211 has been reported to sensitize the dog heart to epinephrine induced arrhythmias at concentrations as low as 2% (v/v) of the inspired air (Beck, *et al.*, 1973).

The information presently available regarding the arrhythmogenic or sensitizing activity of 1211 and related compounds has been derived primarily from whole animal studies based on ECG findings. The experiments reported here are an attempt to determine the direct cardiac mechanism of

of 1211 arrhythmogenesis. By investigating the effects of this agent on cellular transmembrane electrical activity underlying alterations in the spread of excitation in the heart, a basis for arrhythmia production might be suggested. The effects of 1211 alone and in the presence of isoproterenol, a beta adrenergic agonist, on canine Purkinje electrical activity are presented.

METHODS

Mongrel dogs of either sex were anesthetized with sodium pentobarbital (30 mg/kg). The chests were opened through a lateral thoracotomy and the hearts were rapidly removed and placed in cold Tyrode's solution containing 137.0 mM NaCl, 3.0 mM KCl, 2.7 mM $CaCl_2$, 0.5 mM $MgCl_2$, 12.0 mM $NaHCO_3$, 1.8 mM NaH_2PO_4 and 5.5 mM glucose. Strands of Purkinje fibers were dissected free from either ventricle and stored in oxygenated Tyrode's solution until used.

Purkinje fiber preparations were secured to the bottom of an 18 ml tissue bath maintained at 35^oC and perfused with Tyrode's solution at the rate of 100 ml/min. The preparation was stimulated at the rate of 60 beats per minute through bipolar platinum electrodes. The stimulus pulse, delivered through a W-P Instruments Model 305-1 stimulus isolator with a three channel digital pulse generator, was 1 msec in duration and adjusted to 1.5X threshold amplitude.

Transmembrane potentials were recorded with glass microelectrodes

positioned on the Purkinje fiber distal to the stimulating electrodes. The resistance of these electrodes was 20-40 meghoms when filled with 2.5 M KCl. Microelectrodes were connected in series with a Ag-$AgCl_2$ electrode holder, a W-P Instruments Model 750 Micro Probe amplifier and one channel of a Tektronix storage oscilloscope. A second Ag-$AgCl_2$ reference cell, in contact with the perfusion medium in the tissue bath, was connected in series with a voltage-duration-ramp generator calibration unit and completed the ground return of the recording system. Signals from the microelectrode were amplified and differentiated and recorded with a Honeywell 1508 multichannel oscillograph.

Following a 30-60 minute period of equilibration, normal characteristics were recorded during a 30 minute control period. The microelectrode recording sites, either both on a Purkinje fiber, or the proximal on the fiber and the distal impaled in a ventricular muscle cell, were not changed once suitable penetrations had been made. The experiment was terminated if a penetration could not be maintained. Calibration for voltage, duration and rate of depolarization were made during this time and again at the end of the experiment.

The following measurements were made on action potentials during the control period and subsequent test periods: Resting potential (RP), measured from zero reference potential to the level at which the action potential was initiated; action potential (AP) amplitude, measured from RP to the peak of the overshoot; AP duration at 50% repolarization (APD_{50}); AP duration at 90% repolarization (APD_{90}); Vmax, maximum rate of upstroke

of the AP; effective refractory period (ERP), measured as the shortest interval at which a conducted response could be elicited, following the insertion of a test stimulus after every seventh drive stimulus; conduction velocity as the rate of conduction between recording sites; threshold voltage of stimuli.

In the series of dose response experiments, 1211 was introduced into the perfusion reservoir at flow rates of 20, 40, 70 and 115 ml/min, producing mean bath concentrations of 48.1, 97.6, 145.6 and 204.7 μg/ml respectively. Following the control period, 1211 was perfused for a period of 25 minutes at each flow rate. Action potential measurements were made at the end of the control period and after each 25 minute exposure. In addition to these measurements, bath samples were taken for pO_2, pCO_2, pH and 1211 measurements at the end of each test period. The tissue bath concentration of 1211 was measured using an automated gas chromatographic head-space analysis. Duplicate one milliliter bath samples were injected into sealed 20 ml glass vials. The vials were placed in a Perkin-Elmer Multifract F40 gas chromatograph and allowed to equilibrate at 36°C for thirty minutes. The chromatograph utilized a flame ionization detector and a stainless steel column 2 M x 2.1 mm ID packed with Chromasorb 102 (80-100 mesh). The carrier gas was purified nitrogen set at a flow rate of 50 ml/minute. The detector was maintained at 220°C, the column at 180°C, the injector at 210°C and the automated injection needle at 130°C. A duplicate set of four standards

was prepared each day by injecting 1.0 ml of normal Tyrode's solution and 25.0 to 1000.0 μl of an 8% (v/v) 1211-air mixture into sealed vials. The standards were allowed to equilibrate for 60 minutes before being placed in the Multifract F40 with the tissue bath samples. The chromatograph was programmed to perform the following analysis steps: A one second injection, a two minute analysis, a twelve second flushout, a twelve second stabilization period and a single analysis per vial. The known 1211 concentrations and corresponding chromatogram peak heights of the standards were used to prepare a standard curve using linear regression analysis. With the standard curve and the density of 1211 vapor (6.90 μg/ml for 100% 1211 vapor at 25°C and 1 atmosphere), the bath sample peak heights were converted to μg/ml concentrations.

The following protocol was used for the experiments with 1211 and isoproterenol. Following the control period, sodium EDTA (5×10^{-5} M) was added to the perfusion medium to prevent the degradation of isoproterenol (Rosen *et al.*, 1977). The effect of EDTA alone was determined. Isoproterenol (Winthrop Laboratories) was then added to the 500 ml perfusion reservoir by serially diluting a 5×10^{-5}M stock solution to give final bath concentrations of 10^{-9}, 10^{-8} and 10^{-7} M. The effects of each concentration was determined after ten minutes contact. Following the total thirty minute exposure to isoproterenol, a thirty minute washout period was performed, with a second set of control tests made. 1211 was then introduced at a flow rate of 11 ml/min., producing an average bath concentration of 24.5 μg/ml. The isoproterenol series was then repeated.

The data presented here is analyzed by comparison of means using Students T-test. (This same data will be analyzed for future submission by analysis of covariance to take into account the variation in bath concentration at different flow rates).

RESULTS

Range Finding Experiments - An initial set of experiments was performed to (1) determine the effective range of 1211 concentrations which produced measurable responses in the parameters in questions, and (2) determine the time of onset and duration of effects. A flow of 11 ml/min of 1211 alone produced no measurable changes from control. The threshold effect was produced by a flow of 20 ml/min while the maximum response we could measure was seen at a flow 115 ml/min. (This was limited by the changes in pO_2 and pH produced by flows above 115 ml/min). It was determined that bath concentration reached a stable maximum value of 1211 for any given flow rate at 15 minutes following the initiation of 1211 flow. Also, the changes in AP parameters were maximum by 17 or 18 minutes and remained unchanged as long as flow of 1211 was maintained. Twenty-five minutes was selected as the period of exposure at each flow. In several experiments, it was determined that the effects of 1211 were readily reversible in that the tissue characteristics returned to control values with 15-30 minutes of washout.

Nitrogen Control Experiments - To assure that the introduction of 1211 would not compromise the partial pressure of oxygen or alter the pH, five

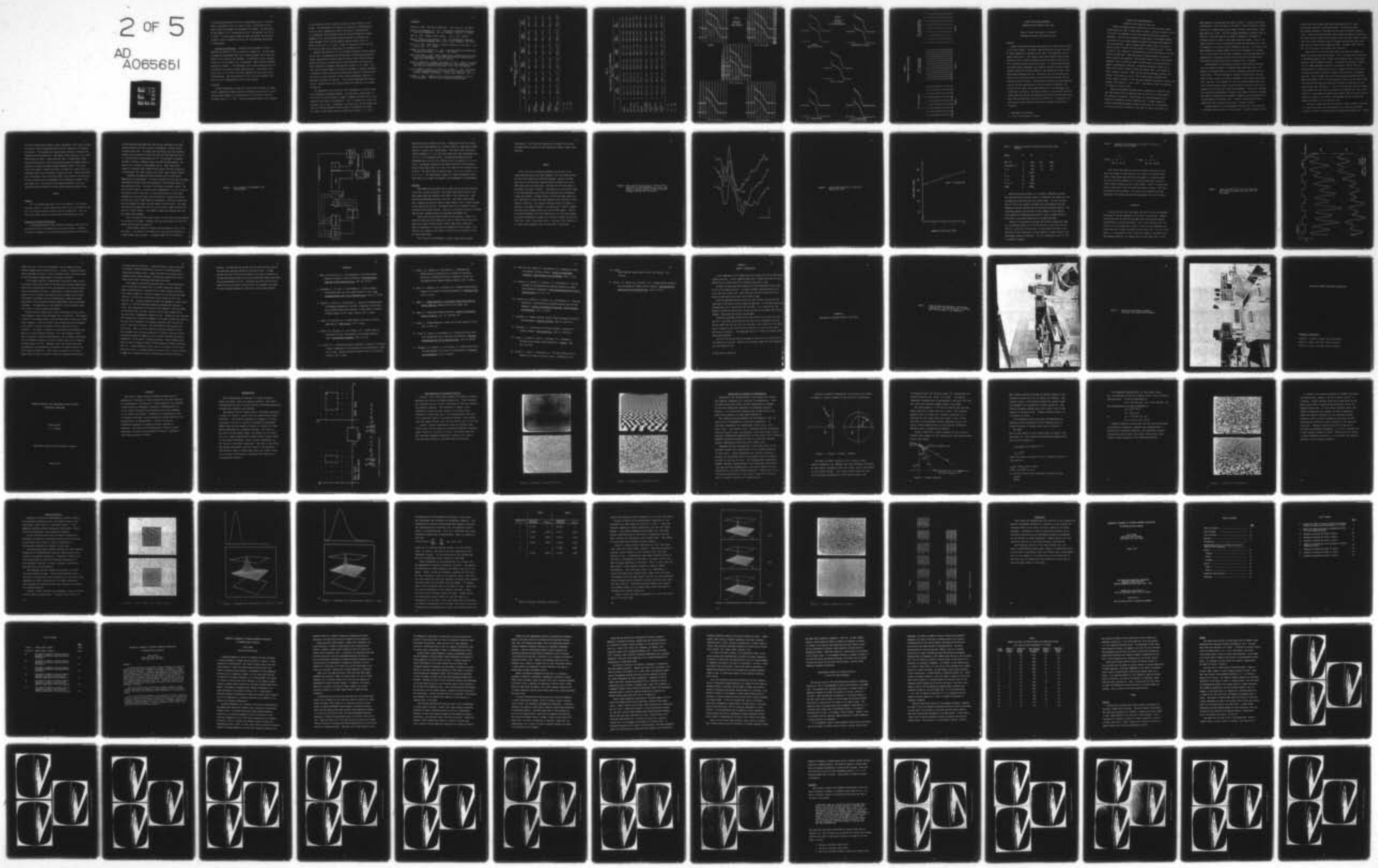
experiments were performed in which nitrogen was substituted for 1211 at identical flow rates. In these experiments, there were no significant changes ($p = 0.5$) in any parameters provided pO_2 was maintained above 200 mmHg. Likewise, pH changes, up to 7.45, had no effect on these parameters. In the experiments with 1211 at a maximum flow of 115 ml/min, pO_2 never fell below 250 mmHg and pH never exceeded 7.45.

1211 Dose Response Experiments - The oscillograph records of one experiment are presented in Figure 1. Depicted are the control and 1211 flows at 20, 40, 70 and 115 ml/min. In this experiment, both recording sites were on the Purkinje fiber. The proximal AP appears above the distal AP in each frame. The effect on both the APD_{50} and APD_{90} may be seen. The other AP characteristics, including APDs, are presented in Table I. These values are the means ± S.E. for the control and each flow. The values for conduction are not presented, but this parameter was noted to increase with each successive concentration of 1211. Stimulation threshold generally decreased with increasing 1211 concentrations.

1211 Plus Isoproterenol - Experiments with 1211 and isoproterenol were conducted to determine whether 1211 sensitizes the Purkinje fiber to adrenergic agents. The results of these experiments are presented in Figure 2 and Table II. The first three frames of Figure 2 are nearly identical. When the effects of EDTA are compared to the control (Table II), one can see that EDTA produces essentially no effect of its own. The second frame

AD-A065 651 OHIO STATE UNIV RESEARCH FOUNDATION COLUMBUS F/G 1/3
USAF-ASEE (1978) SUMMER FACULTY RESEARCH PROGRAM (WPAFB). VOLUM--ETC(U)
NOV 78 C D BAILEY F44620-76-C-0052
UNCLASSIFIED AFOSR-TR-79-0232 NL

2 OF 5
AD
A065651



is following the washout of the initial isoproterenol series. The third frame is the effect of 1211 at a flow 11 ml/min., indicating that this flow is below the threshold for measurable 1211 effects. The bottom two are the effects of 10^{-7} isoproterenol and 10^{-7} isoproterenol plus 1211 at 11 ml/min. It can be seen by comparing these two records and the values in Table II that a noneffective flow of 1211 can potentiate the effects of isoproterenol

Automaticity Experiments - Experiments were performed in order to determine the effects of 1211 on ventricular automaticity. We were unable to produce an experimental model with a steady pacemaker rate to test the effects of 1211 alone on this parameter. We were able to show a potentiation by 1211 on the effects of an isoproterenol induced Purkinje pacemaker preparation (Figure 3). The spontaneous rate of this preparation with 10^{-7} isoproterenol was stable at twenty impulses/minute. Ten minutes of perfusion with 1211 at a flow of 40 ml/min enhanced the rate to over thirty impulses/minute. Upon discontinuation of 1211, the rate returned to the original value of twenty/minute within fifteen minutes.

DISCUSSION

In these experiments, we were able to show that 1211 alone, at concentrations comparable to those attained by breathing less than 5% 1211, could produce marked changes in AP characteristics and ventricular electrical parameters (Beck *et al.*, 1973). The most pronounced effect of this compound

is its shortening of the AP duration, especially that of Phase 2 or the plateau. The shortening of the plateau is significant at concentrations of only 48 μg/ml, progressing to 55% of the control at maximum levels. It is noteworthy that the plateau phase has been attributed to an inward movement of Ca++ and the changes seen here may well be related to the inward movement of Ca++ and the negative inotropic effect of 1211 in muscle tissue (Toy *et al.*, 1976). Though not reported, we found that the AP of muscle fibers also shorten under the influence of 1211.

Regarding the effects of 1211 on arrhythmia production, several factors may be cited. By decreasing the refractoriness in the Purkinje conducting system, impulses which normally could not re-enter their original pathway because of the usual refractoriness following conduction could now do so under conditions of decreased refractoriness (Hoffman *et al.*, 1975). This could lead the establishment of re-entrant arrhythmias. Secondly, decreases in conduction velocity could also be deleterious, especially from the standpoint that an area of the heart may be excited retrogradely, setting up additional conditions for re-entry. That threshold decreased with increasing 1211 concentrations could provide evidence for conditions of increased excitability.

The experimental results with 1211 and isoproterenol are equally interesting regarding arrhythmia production. We found that a concentration of 1211, which by itself produces no changes in membrane characteristics, could potentiate the effects of isoproterenol. This is evidence for a direct sensitizing action of 1211. Furthermore, the finding that 1211 may enhance the rate of a Purkinje pacemaker provides support for production of ectopic foci shown in whole animal studies on sensitization (Beck *et al.*, 1973).

REFERENCES

Aviado, D., 1974. Toxicity of propellants. *Prog. Drug Res.,* 18: 365-97.

Back, K.C. and VanStee, E.W., 1977. Toxicology of haloalkane propellants and fire extinguishants. *Ann. Rev. Pharmacol. Toxicol.,* 17: 83-95.

Bass, M., 1970. Sudden sniffing deaths. *J.A.M.A.*, 212: 2075-79.

Beck, P.; Clark, D. and Tinston, D., 1973. The pharmacologic actions of bromochlordifluormethane (BCF). *Toxicol. Appl. Pharmacol.*, 24: 20-29.

Harris, W., 1973. Toxic effects of aerosol propellants on the heart. *Arch. Intern. Med.*, 131: 162-66.

Hoffman, B; Rosen, M and Wit, A., 1975. Electrophysiology and pharmacology of cardiac arrhythmias. *Am. Heart J.*, 89: 115-22.

Levy, A and Lewis T., 1911. Heart irregularities resulting from the inhalation of low percentages of chloroform vapor and their relationship to ventricular fibrillation. *Heart*, 3: 99-112.

Rosen, M., Hordorf, A., Ilvento, J and Danilo, P. 1977. Effects of adrenergic amines on electrophysiological properties and automaticity of neonatal and adult canine Purkinje fibers. *Circ. Res.*, 40: 390-400.

Toy, P, VanStee, E, Harris, A., Horton, M. and Back, K., 1976. The effects of three halogenated alkanes on excitation and contraction in the isolated, perfused rabbit heart. *Toxicol. Appl. Pharmacol.*, 38: 7-17.

VanStee, E., 1974. A review of the toxicology of halogenated fire extinguishing agents. AMRL-TR-74-143, Wright-Patterson AFB, Ohio

TABLE I

CANINE PURKINJE AP - 1211 DOSE RESPONSE
(mean ± S.E.)

	N	RP (mVolts)	AP (mVolts)	APD_{50} (msec)	APD_{90} (msec)	Vmax (V/sec)	ERP (msec)
Control	8	95.2 ± 0.9	139.2 ± 1.6	266.3 ± 12.7	366.5 ± 13.8	494.5 ± 18.2	332.9 ± 18.7
Flow @ 20.0 ml/min	6	94.7 ± 0.4	136.7 ± 1.3	232.6 ± 18.1*	337.1 ± 19.6	484.0 ± 21.8	307.3 ± 19.3
Flow @ 40.0 ml/min	8	91.4 ± 1.5*	134.0 ± 0.8+	194.5 ± 11.8+	327.4 ± 15.6	463.1 ± 15.0	298.9 ± 15.1
Flow @ 70.0 ml/min	8	91.3 ± 1.2†	130.9 ± 1.4+	169.6 ± 10.2+	314.3 ± 12.3†	447.3 ± 11.8*	282.5 ± 10.1*
Flow @ 115.0 ml/min	8	89.5 ± 2.1†	127.8 ± 2.7+	147.3 ± 8.6+	305.0 ± 10.7+	428.4 ± 15.6†	278.0 ± 8.8†

* $p < .05$

† $p < .01$

+ $p < .005$

TABLE II

CANINE PURKINJE AP - 1211 PLUS ISOPROTERENOL
(mean ± S.E.)

	N	RP (mVolts)	AP (mVolts)	ADP_{50} (msec)	ADP_{90} (msec)	Vmax (V/sec)	ERP (msec)
Control	6	93.1 ± 1.4	129.1 ± 0.9	231.0 ± 16.1	311.3 ± 12.6	421.2 ± 26.6	281.0 ± 11.2
EDTA	6	93.7 ± 1.5	129.7 ± 1.0	231.8 ± 15.4	312.1 ± 12.2	424.7 ± 29.0	279.5 ± 12.7
ISO 10^{-9} M	6	93.7 ± 1.5	130.0 ± 1.0	230.3 ± 15.7	309.8 ± 12.1	424.7 ± 29.0	278.0 ± 12.3
ISO 10^{-8} M	6	94.0 ± 1.6	131.0 ± 1.2	210.6 ± 14.0	285.6 ± 9.9	436.0 ± 28.0	258.4 ± 10.3
ISO 10^{-7} M	6	94.5 ± 1.4	131.3 ± 1.1	194.7 ± 10.5	258.3 ± 8.6†	439.7 ± 27.1	235.6 ± 7.9†
Washout	6	94.3 ± 1.8	129.9 ± 1.0	236.2 ± 16.9	319.6 ± 13.3	414.8 ± 26.3	285.6 ± 12.0
1211 Flow @ 11.0 ml/min.	6	92.9 ± 1.6	126.7 ± 0.9	228.0 ± 15.4	322.0 ± 12.9	399.5 ± 28.1	286.4 ± 11.6
ISO 10^{-9} M + 1211	6	93.1 ± 1.3	126.3 ± 1.6	217.4 ± 17.2	318.1 ± 12.1	416.3 ± 26.1	283.2 ± 9.2
ISO 10^{-8} M + 1211	6	93.5 ± 2.2	128.3 ± 1.2	198.5 ± 13.0	287.1 ± 7.3	423.8 ± 18.7	255.3 ± 7.2*
ISO 10^{-7} M + 1211	6	93.4 ± 2.1	128.1 ± 1.3	174.2 ± 9.4†	248.5 ± 6.3+	429.3 ± 19.8	222.0 ± 5.3+

* $p < .05$

† $p < .01$

+ $p < .005$

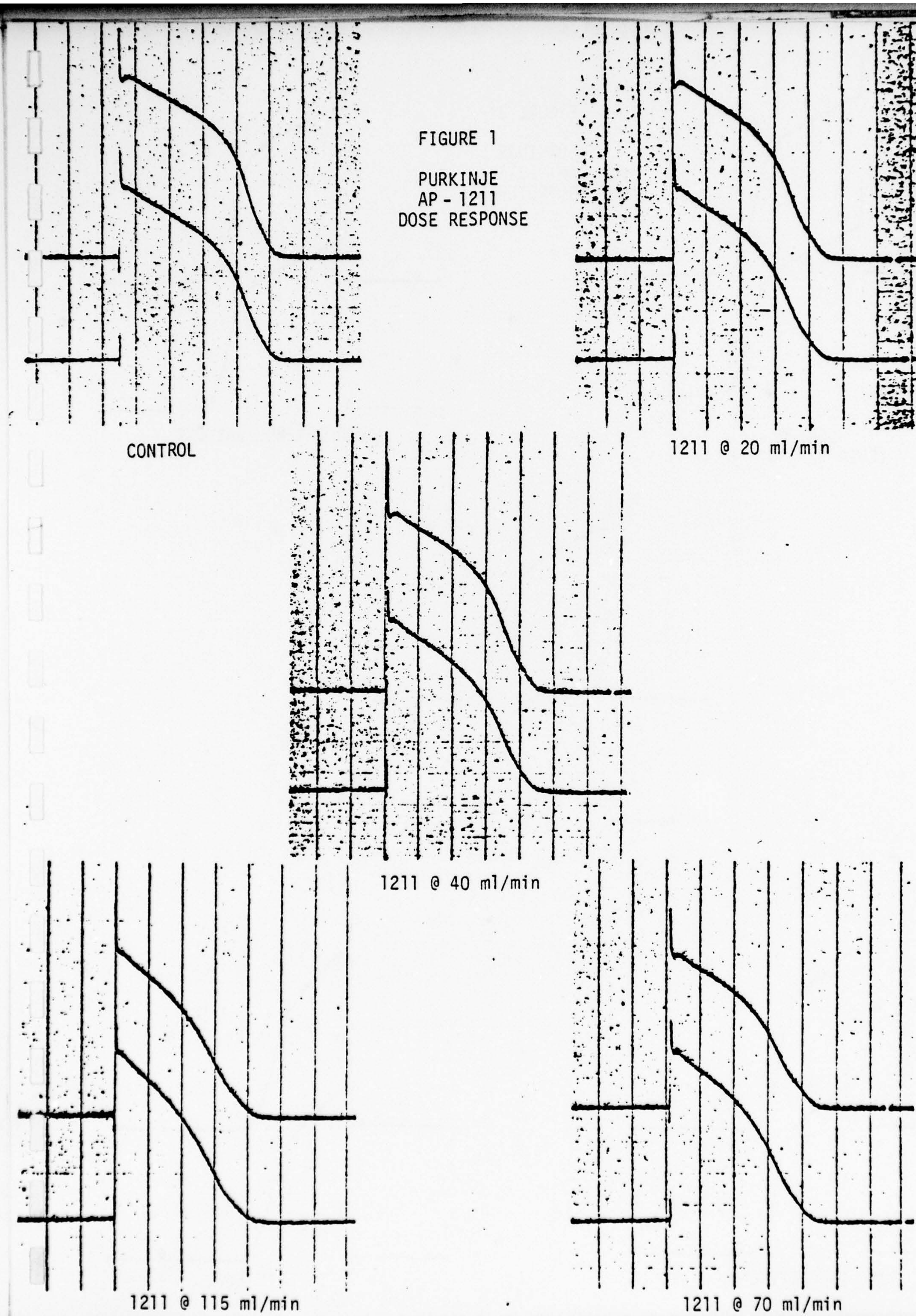


FIGURE 1

PURKINJE
AP - 1211
DOSE RESPONSE

FIGURE 2

PURKINJE
AP - 1211
PLUS ISOPROTERENOL

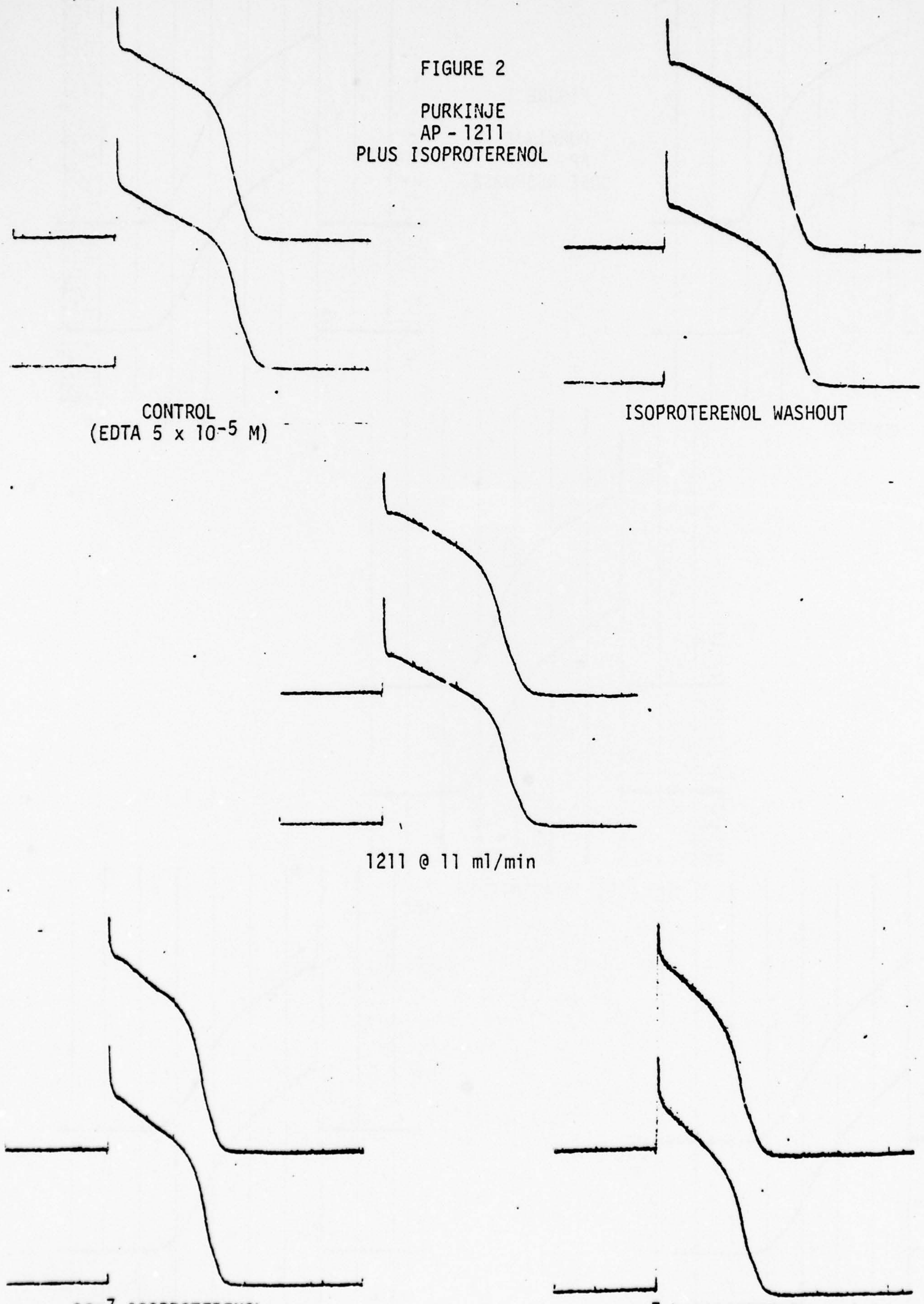
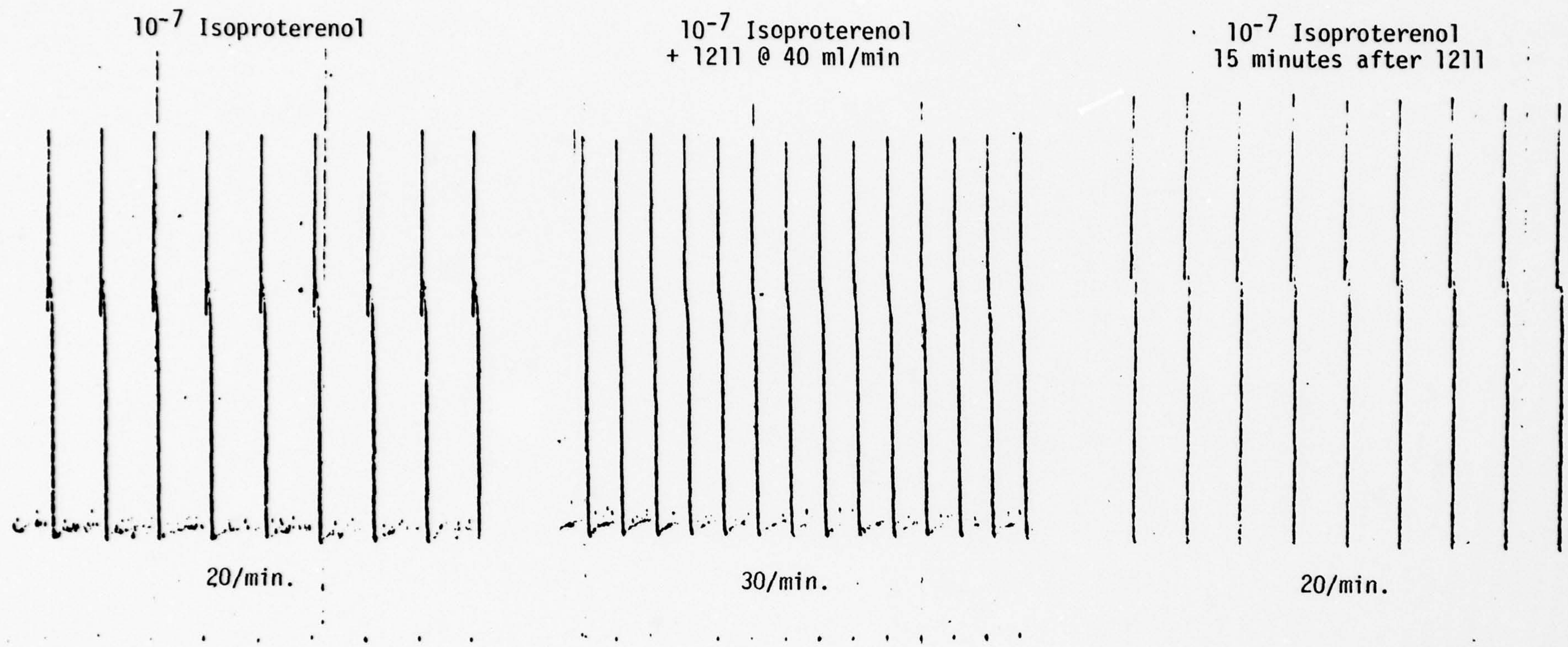


FIGURE 3

SPONTANEOUS PURKINJE PACEMAKER



STEADY STATE EVOKED RESPONSES: INTERACTION WITH COGNITIVE TASK LOAD

Glenn F. Wilson* and Robert D. O'Donnell**

Wittenberg University, Springfield, Ohio

Abstract

Steady state evoked responses have been used to study several aspects of the visual system. The present study was designed to assess the value of these responses as a measure of work load or task complexity using a memory-scanning task. Even though steady state responses have not been previously used in cognitive research it was felt that they may be useful in monitoring cognitive work load levels. If usable these responses would have several advantages over currently available transient evoked response methods of assessing attention, task complexity, and work level. Among these advantages would be: noninvasive into to the primary task, they would not require diversion of resources from the primary task and acquisition would be faster. Data has been collected but evoked responses have not yet been obtained due to difficulty acquiring computer time. Analysis of the steady state evoked responses will be performed as soon as the data is available. Preliminary analysis of the available transient evoked responses show that the paradigm used elicited reliable memory-scan data which provided three levels of task difficulty. Discrepancies with one transient evoked response study from the relevant literature are discussed.

* Department of Psychology

** Lt Col, United States Air Force

Steady State Evoked Responses:

Interaction with Cognitive Task Load

Recent electrophysiological studies have utilized transient average evoked potentials (AEP) to study cognitive events. Using this method individual stimuli are presented to the S which require some type of cognitive processing. The S's brain responses to these stimuli are averaged together and correlations are sought between the cognitive processes and segments of the associated AEPs. Several seconds elapse between stimuli allowing the nervous system to recover from the effects of the preceding event. The AEP elicited in this situation is characterized by several deflections of variable latency and amplitude. These amplitude and latency measures are typically used to quantify the responses. One reliable finding is that a late positive component, the so called P3 or P300, has been found to be associated with a number of psychological processes: uncertainty resolution (Sutton, et al, 1965), discrimination (Ritter, Simson, and Vaughan, 1972), task complexity (Poon, Thompson, and Marsh, 1976), decision making (Rohrbaugh, Donchin, and Eriksen, 1974; Squires, N. et al, 1977), discrimination problem solving (Wilson, Harter, and Wells, 1973), and task relevance (Courchesne, Hillyard, and Galambos, 1975; Squires, K. et al, 1977). See Donchin, Ritter, and McCallum (in press) for current review.

Studies investigating variables such as attention,task difficulty and work load operate with the assumption that the brain has a limited capacity for processing information. These channels of processing capability are parcelled out among the various competing tasks. A common approach to studying the electrophysiology of these attention and work load processes is to vary the focus of attention or the level of load on the subject and

assess changes in the amplitude and latency of P300. A finding of particular interest here is that the latency of the P300 is directly related to reaction times (RT) in a memory-scanning task (Gomer, Spicuzza, and O'Donnell, 1976; Adam and Collins, 1978). Using the paradigm developed by Sternberg (1969 (a), 1969 (b)) it was found that as the number of elements in the memory set increased the RTs to identify these items likewise increased. That is, the larger the set of items to be scanned the longer the time necessary to identify whether or not a given stimulus belonged to that set. The latency of the P300 component also became longer as the size of the memory set increased. This is a very useful task for those interested in the electrophysiological concomitants of increased processing load since both RT and P300 latency show linear functions as the size of the memory set is increased.

A second method of investigating attention, task difficulty and work load; one especially useful in situations where the primary task is not amenable to transient AEPs, is to utilize a secondary task which uses discrete stimuli. AEPs are collected in response to these stimuli and some aspect of them is used as an index of attention or work load of the primary task. Many real life situations fit into this category because most tasks do not involve discrete stimuli that can be used to provide an AEP, this is especially true in the case of work load assessment. The use of a secondary task presents problems because it requires that some of the brain's capacity be assigned to processing this secondary information which may detract from the performance of the primary task. A method of measuring work load is needed which does not possess this actual or potential problem.

Another method of eliciting responses from the brain is by providing a continuous stimulus such as a flicking light. The AEPs collected in this

situation are called steady state AEPs and characteristically appear as sinusoidal waveforms of uniform amplitude. The amplitude, phase and frequency features of these AEPs are the typically measured characteristics. Steady state AEPs have been used to study the mechanisms underlying the functioning of complex sensory systems such as the visual system. Spectral sensitivity, depth perception; spatial frequency and many more topics have been areas of research using steady state AEPs. See Regan (1972, 1977 (a), 1977 (b)) for reviews of the steady state literature.

Regan (1977 (b)) has stated that steady state AEPs are probably not useful for studying cognitive processes. This is no doubt due to the nature of the methods needed to elicit steady state responses, that is, one which requires a rapidly repeating stimulus. The present investigation, however, will use steady state AEPs elicited by a background stimulus while the subjects perform a primary cognitive task. Steady state AEPs will be used to assess the level of load that a subject is working under in the memory-scan task. Since the human brain has limited processing capacity(Shiffrin, McKay, and Shaffer, 1976) increased resource utilization by the primary cognitive task should decrease the resources available for processing the background steady state information. If this is the case there should be changes in the steady state AEPs which are related to the subjects work load or task difficulty. Since cognitive effects upon transient AEPs are found in central regions of the brain, this area will be monitored for both transient and steady state AEPs as will the occipital area.

The use of steady state AEPs to assess the level of subject involvement in a primary task would have the advantage of not requiring the subject to perform a possibly interferring and/or confounding secondary task. Studies

utilizing transient AEPs (Wickens, Isreal, and Donchin, 1977; Isreal, Wickens, and Donchin, 1978) have measured P300 of the AEP response to the secondary task stimuli. The secondary task required some diversion of attention and resources from the primary task. Additionally, they found that "first order" effects were not useful in monitoring work load. The amplitude of P300 "declined precipitously as soon as the tracking task was imposed" making it necessary to analyze the target stimulus sequence effects. As the work load on the subject increased the "reach" or effect of a target stimulus on subsequent target stimuli decreased as measured by P300. Steady state AEPs, due to the nature of their eliciting stimuli, do not require the diversion of one's attention with the subsequent loss of processing "channels" from the primary task. The analyses of the steady state AEPs is also more straight forward and requires less data thereby reducing evaluation time.

METHOD

Subjects

Six right handed young adults with an age range of 17 to 25 years, served as subjects. The paid subjects had normal (5 Ss) or corrected to normal (1S) vision and were not currently taking any medication. They also fell within normal limits for lateral phoria and stereoscopic vision.

Recording and Stimulating Equipment

Electroencephalographic (EEG) recordings were made from central (CZ) and occipital (OZ) sites referenced to the mastoid process. Electrode resistance as measured by a Grass EZM Capacitance meter was below 5k ohms.

An electrode positioned above the right eye was referenced to the same mastoid electrode and used to monitor eye movements. Beckman Ag Ag/Cl electrodes were used. The signals were fed through high input impedance probes into Grass model P511 amplifiers with half amplitude band pass of 0.1 and 100 HZ and an amplification of 10^5. A Grass model 6 polygraph was used to provide a permanent copy of the EEG and event markers. See Figure 1 for a diagram of the equipment set up. These signals were stored on a Honeywell model 5600B FM tape recorder for later averaging on a Nicolet model 1072 signal averager and a Xerox Sigma 5 digital computer.

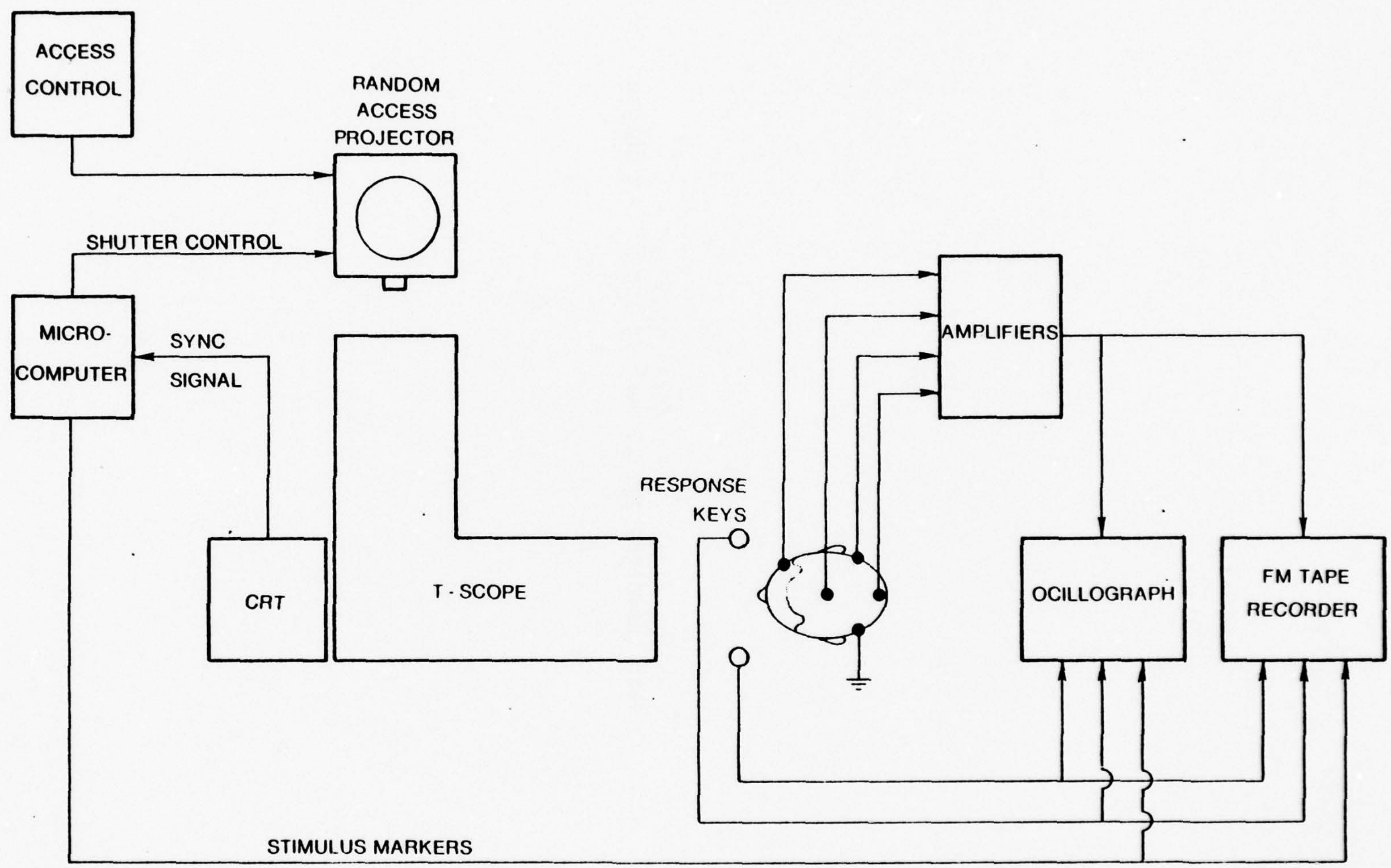
Transient and steady state AEPs consisted of 512 or 1024 points sampled one millisecond apart. In order to separate the transient and steady state AEPs the letter stimuli were presented at four different phases of the background flicker. The onset of the letters occurred at either: the onset of the flicker, or at points which corresponded to 1/4, 1/2, or 3/4 of the period of each flicker rate. The one second duration of the letter presentation, the above timing, and the generation of appropriate markers was done by an E and L model MMD-1/MI microcomputer. When the transient AEP data was averaged the steady state AEP tended to cancel itself. The steady state AEPs were triggered by the first positive going trigger following the onset of a letter stimulus. This tended to remove the transient AEPs from the steady state responses.

RTs to the positive and negative stimuli in each condition were measured to the nearest millisecond. Responses less than 200 or more than 1500 milliseconds were excluded from analysis.

A three channel Scientific Prototype tachistoscope was used to align the stimuli. The letters for the memory-scan task were rear projected by a Kodak random access projector. A Lafayette model 43011-16 electronic

FIGURE 1 Block diagram of the equipment used in this study

DIAGRAM OF EQUIPMENT



shutter was used to present the stimuli. Background flicker and a square fixation point were presented by a Tektronix model 632 video monitor (Gomer and Bish, in press) via a second channel. The letter stimuli were white letters subtending 1° 4" x 0° 56" visual angle which were superimposed upon a 2° 30" x 5° 44" background field. The unpatterned background field alternated from on to off, with a 50% duty cycle, at rates of 3.5, 5.0, or 7.0 HZ. The average luminance of the letters was 167 fl while the background flicker when on was 7.8 fl and 0.14 fl when off for a contrast ratio of 96.5%. The letters used for each set were: H; K, B, V, Z; and D, I, L, M, S, T, W, Y. The interstimulus interval for letter presentation varied from three to six seconds. See Appendix B for photographs of the equipment.

Procedure

After memorizing the three sets of letters the Ss were given practice trials prior to data collection. The Ss were instructed to fixate on the small square which was located at the center of the viewing field. Head position was maintained by using a chin rest. Each index finger rested over a response key which was used to signal whether or not a letter belonged to the current positive set. The subjects were told to respond as quickly as possible while being accurate. They sat in an Industrial Acoustics shielded room and were earphones during the experiment (See Appendix A).

Two-two hour sessions were required for each subject, a total of 15 conditions were used, three memory set levels by three flicker rates, each memory set without flicker and each flicker rate without a memory set. Order of presentation of conditions was randomized for each subject. Each condition was composed of 128 stimuli, 64 positive and 64 negative in one of three random orders.

Trials containing eye movement or other artifacts were excluded

from analysis. The flicker was observed by the subject for at least 30 seconds before the start of each condition to insure a steady state condition.

RESULTS

Due to difficulty in obtaining computer services most of the planned analyses have not yet been completed. Only the transient AEPs for the positive memory set items are available. Analysis of these data show that the previously reported effects of memory set size on P300 latency have been replicated. Transient AEPs from the vertex of one subject are shown in Figure 2. These AEPs are to the positive items of the indicated memory sets during the conditions when no background flashes were present. The mean latency of P300 from these records was 460.5 msec which is within the range reported in the literature of from 250 msec to 500 msec. The latency of P300 was found to increase as a function of the number of letters in the memory set. Figure 3 presents the relationship between positive set size and P300 latency. Analysis of variance performed on the P300 latencies for all flash rates showed significant differences in latency as a function of memory set size and flash rate. Table I shows these results. Analyses of the latencies of all other vertex components were not statistically significant.

FIGURE 2 Vertex AEPs from one subject in the no flicker conditions for the three memory set sizes. Negativity is plotted upwards, calibration markers represent 200 msec and 5 microvolts.

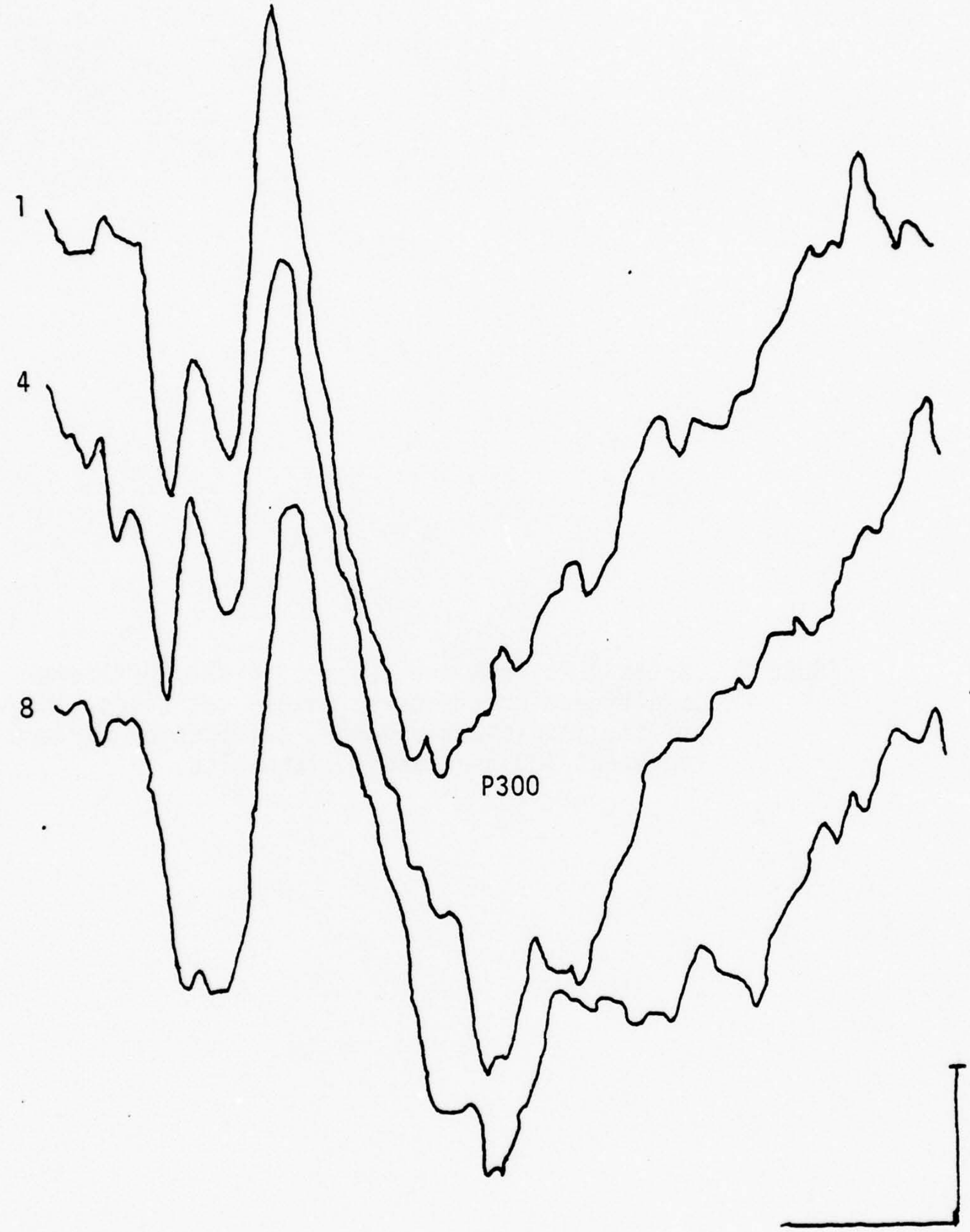

1
4
8
P300

FIGURE 3 Vertex P300 latencies as a function of positive set size.

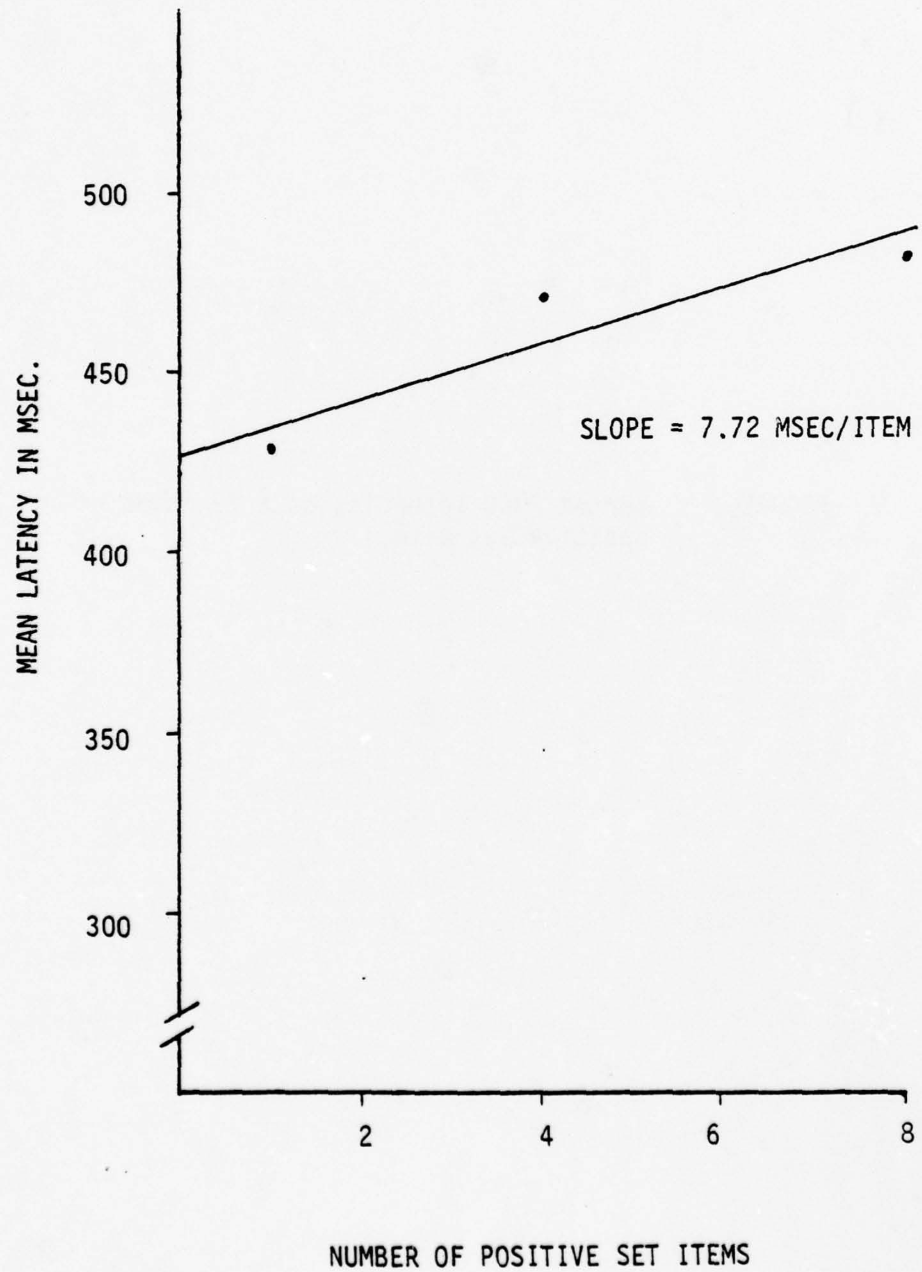

500
450
400
350
300
MEAN LATENCY IN MSEC.
SLOPE = 7.72 MSEC/ITEM
2
4
6
8
NUMBER OF POSITIVE SET ITEMS

Table I Results of Analyses of Variance on the Vertex P300 Latencies

SOURCE	df	SS	F	P
MSET (M)	2	41129	20.5	.0005
FLASH RATE (F)	3	6444	3.5	.041
F x M	6	2919	.40	.874
S	5	125637		
S x M	10	10033		
S x F	15	9201		
S x M x F	30	36524		
TOTAL	71	231887		

While RTs are not available it is certainly reasonable to assume that they would show the usual result of increasing as the memory set size increases thus replicating this very robust effect. The very low error rate of 0.9% for our subjects indicates that the procedures used were adequate to provide reliable data from the memory-scanning paradigm. Thus, the memory-scanning task provided reliable changes in P300 latencies which support the literature and establish a task of graded difficulty with which to test these effects upon the steady state AEPs.

Analysis of the occipital transient AEP latencies showed that there was a significant effect on P1 ($F=3.98$, $df=3/15$, $p < .03$) and P2 ($F=4.05$, $df=3/15$, $p < .03$) due to flicker rate. As can be seen from Table 2 the condition in which there was no flicker resulted in longer latency P1 and and shorter latency P2 components. RTs will be examined to test for similar behavioral changes.

Table II Latencies in Milliseconds of Occipital P1 and P2 as a Function of Flicker Rates

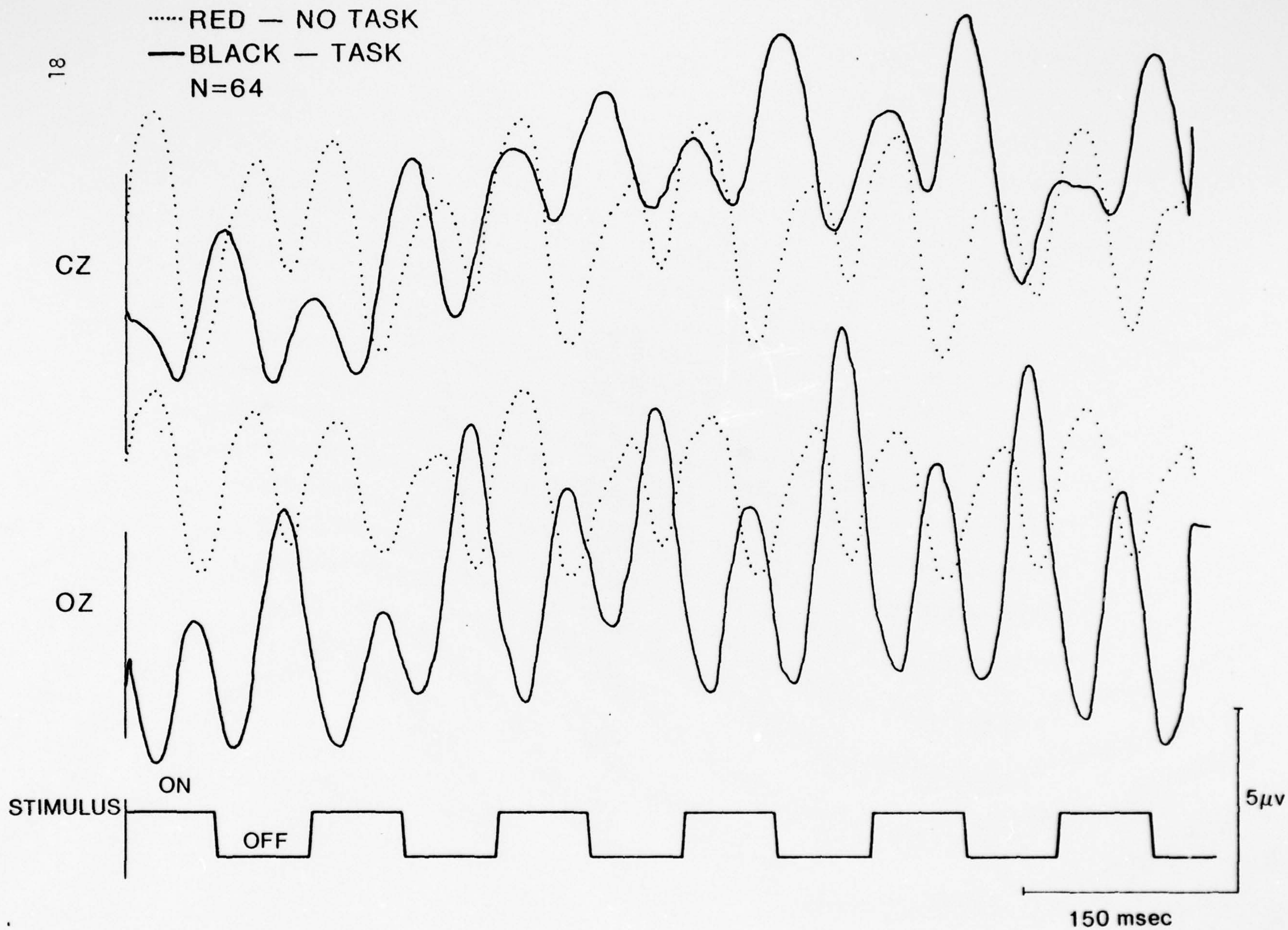
P1					P2				
Flicker	0	3.5	5	7	Flicker	0	3.5	5	7
	108	91	94	90		218	250	252	252

While steady state AEPs are not yet available, pilot data are suggestive of changes in these potentials as a result of cognitive factors. Figure 4 shows steady state AEPs from one subject in a condition where only the flicker was present and in a condition where both flicker and letter stimuli of memory set size 4 were present. There is a marked phase shift of these AEPs in both the vertex and occipital records. Amplitude differences are also apparent. This data is only suggestive but does demonstrate that steady state AEPs were generated in our test situation.

DISCUSSION

Since the critical data, the steady state AEPs, are not yet averaged, discussion of the main hypothesis of the study is not possible. However, the available data do show that the memory scan paradigm used produced graded effects in vertex P300 latency. This is very important since the validity of this paradigm must be shown in order to allow statements concerning the use of steady state AEPs to monitor work load. With this information it is now possible to see if there are changes in the steady state AEPs between conditions of no memory scan task and graded levels of the

FIGURE 4 Steady state AEPs for flicker only (No Task) and flicker plus memory scan task (Task) for one subject.



······RED — NO TASK
—BLACK — TASK
N=64
CZ
OZ
STIMULUS
ON
OFF
5μv
150 msec

memory scan task. This will show whether or not the steady state AEP reflects changes due to cognitive activity. Further, if there are effects due to the memory scan task, it will be possible to see if the three levels of the task produce graded changes in the steady state AEPs.

If the steady state AEP is demonstrated to provide an index of cognitive activity many possibilities for application and further research exist. The relative noninvasiveness of this technique make it an excellent candidate for operational utilization. Additionally, it does not require the subject to divert attention and resources from the primary task or tasks. This method of assessment also has the advantage of speed since steady state AEPs can be collected very rapidly which permits quick and continual assessment. The speed aspect is very important in many situations, especially those in rapidly changing environments.

Further areas of research would involve the testing of the so called "high frequency" steady state AEP (Regan, 1977, (a) and (b)). This response is found in the frequency range of 50 to 60 Hertz which is above the flicker fusion frequency. Since the flicker of the stimulus would not be apparent to the subject it would not interfer with the primary task in most situations. This factor, in addition to those previously mentioned, would make this a highly desirable technique. This "high frequency" steady state is reported to have different anatomical loci, different color properties and different responses to stimulus intensity than the low frequency response (Regan, 1977 (a)). Therefore, even if the present study does not demonstrate the hypothesized relationship with cognitive factors the high frequency response may. Other areas of possible utility for the steady state AEP would be studies of attention, hemisphere lateralization

and other cognitive situations. Positive findings in these areas could be extremely important theoretically as well as in neurology, sensory functioning and human factors. Steady state AEPs are currently used to diagnose nervous system pathology, the described research could lead to studies which would investigate psychological disorders.

With regard to the available transient AEPs it should be mentioned that our data does not support that of a recent study which used the memory scanning paradigm. Adam and Collins (1978) reported significant P300 latency changes as a function of memory set size from many electrode locations. We did find a strong effect at the vertex but not at the occipital site. They also reported two P300 like components in their vertex records, i.e. P270 and P350 whose latencies increased with larger set sizes. They hypothesize that the earlier P270 component, which was hidden by the P350 at small set sizes, reflects a serial search process while the later P350 is independent of memory set size. Inspection of our transient AEPs revealed only one positive component in the range of P300 (See Figure 2). We could find no evidence of two separate deflections even with our set size of ight positive letters. These discrepancies could be due to several factors. Adam and Collins used the varied-set procedure in which they presented their set digits sequencially followed by the test stimulus. We asked the subjects to memorize a list of stimuli and then individually presented all of the stimuli, fixed-set procedure. These different procedures are known to produce slightly different behavioral effects (Sternberg, 1969 (a)). A more problematic issue is that of the timing of their stimuli which could lead to a contingent negative variation (CNV). The test stimulus followed the last memory set digit by 1300 msec and a warning stimulus by

800 msec. The reported time constant of their amplifiers and possibly the electrodes used would permit the recording of CNVs. It seems possible that their P270 could be related to the return to baseline of the CNV which became evident as the task related P300 latency grew longer with increased memory set size. Since Adam and Collins did not have a condition which would permit the testing of this hypothesis the reason for the discrepancy between our results must await further research.

REFERENCES

1. Adam, N. and Collins, G.I., "Late Components of the Visual Evoked Potential to Search in Short-term Memory," Electroencephalography and Clinical Neurophysiology, 1978, 44, 147-156.

2. Courchesne, E., Hillyard, S., and Galambas, R., "Stimulus Novelty, Task Relevance and the Visual Evoked Potential in Man," Electroencephalography and Clinical Neurophysiology, 1975, 39, 131-143.

3. Donchin, E., Ritter, W., and McCallum, W., "Cognitive Psychophysiology: The Endogenous Components of the ERP," Proceedings of the Conference on Event-Related Brain Potentials in Man, National Institute of Mental Health, Airlie House, Virginia, 1977, in press.

4. Gomer, F.E. and Bish, K.G., "Evoked Potential Correlates of Display Image Quality," Human Factors, 1977, in press.

5. Gomer, F.E., Spicuzza, R.F., and O'Donnell, R.D., "Evoked Potential Correlates of Visual Item Recognition during Memory-Scanning Tasks," Physiological Psychology, 1976, 4, 61-65.

6. Hillyard, S.A., "Electrophysiological Assessment of Attentional Processes in Man," Proceedings of the Conference on Event-Related Brain Potentials in Man. National Institute of Mental Health, Airlie House, Virginia, 1977, in press.

7. Isreal, J.B., Wickens, C.D., and Donchin, E., "P300 Amplitude Changes during a Tracking Task as a Function of Continuous Variations in Tracking Difficulty," Presented at Society for Psychophysiological Research Meeting, Madison, 1978, in press.

8. Poon, L.W., Thompson, L.W., and Marsh, G.R., "Average Evoked Potential Changes as a Function of Processing Complexity," Psychophysiology, 1976, 13, 43-49.

9. Regan, D., Evoked Potentials in Psychology, Sensory Physiology and Clinical Medicine, Chapman and Hall Ltd, London, 1972.

10. Regan, D., "Steady-state Evoked Potentials," Journal of the Optical Society of America, 1977, 67, 1475-1489, (a).

11. Regan, D., "Evoked Potentials in Basic and Clinical Research," Paris, 1977, in press (b).

12. Ritter, W., Simson, R., and Vaughan, H.G.,"Association Cortex Potentials and Reaction Time in Auditory Discrimination," Electroencephalography and Clinical Neurophysiology, 1972, 33, 547-555.

13. Rohrbaugh, J.W., Donchin, E., and Eriksen, C.W., "Decision Making and the P300 Component of the Cortical Evoked Response," Perception and Psychophysics, 1974, 15, 368-374.

14. Shiffrin, R.M., McKay, D.P., and Shaffer, W.D., "Attending to Forty-nine Spatial Positions at Once," Journal of Experimental Psychology: Human Perception and Performance, 1976, 2, 14-22.

15. Squires, K.C., Donchin, E., Herning, R.I., and McCarthy, G., "On the Influence of Task Relevance and Stimulus Probability on Event-Related Potential-Components," Electroencephalography and Clinical Neurophysiology, 1977, 42, 1-14.

16. Squires, N.K., Donchin, E., Squires, K.C., and Grassberg, S., "Bisensory Stimulation: Inferring Decision-related Processes from the P300 Component," Journal of Experimental Psychology: Human Perception and Performance, 1977, 3, 2-9-315.

17. Sternberg, S., "Memory-scanning: Mental Processes Revealed by Reaction-time Experiments," American Scientist, 1969, 57, 421-457.(a)

18. Sternberg, S. "The Discovery of Processing Stages: Extensions of Donders' Method," Acta Psychologica, 1969, 30, 276-315.(b)

19. Sutton, S., Braren, M., Zubin, J., and John, E.R., "Information Delivery and the Sensory Evoked Potentials," Science , 1965, 150, 1187-1187.

20. Wickens, C., Isreal, J., and Donchin, E., "The Event-related Cortical Potential as an Index of Task Work Load," Proceedings of the

20. CONT'D
Annual Meeting, Human Factors Society, San Francisco, 1977, in press.

21. Wilson, L.E., Harter, M.R., and Wells, H.H., "Evoked Cortical Potentials and Discrimination Problem-solving in Humans," Electroencephalography and Clinical Neurophysiology, 1973, 34, 15-22.

APPENDIX A

SUBJECT'S INSTRUCTIONS

In this experiment we will measure your brain waves while you are performing a memory-scan task. You will memorize three sets of letters and then pick those letters out of a series which will be shown to you one at a time.

In order to record your brain waves we will place electrodes on your head to pick up the ongoing activity In order to obtain a good recording we will clip a very small area of hair, then clean this area with alcohol and finally attach an electrode to your scalp with a piece of tape.

Once the experiment begins we would like you to press a button with your right hand if the letter is one of those you memorized. If the letter was not one you memorized, press a button with your left hand. It is important that you respond as quickly as possible but also be sure that you are making the correct response. Both speed and accuracy are important.

During the experiment please refrain from blinking until after the letter goes off. Eye blinks during the time the letters are on contaminate the data. Use the square that you will see to fix your gaze. Don't stare at it but rathe keep it centered in your field of view so that you will not miss the letters when they are flashed on. The background light will flicker at different rates during the experiment.

We will first let you practice the task so that you will be familiar with the conditions to be used. Then we will run several conditions, each separated by a few minutes of rest.

Do you have any questions?

APPENDIX B

Photographs of apparatus used in this study.

Photo 1. Stimulus generating equipment; tachistoscope, video monitor, and slide projectors. The Grass amplifiers are shown in the background.

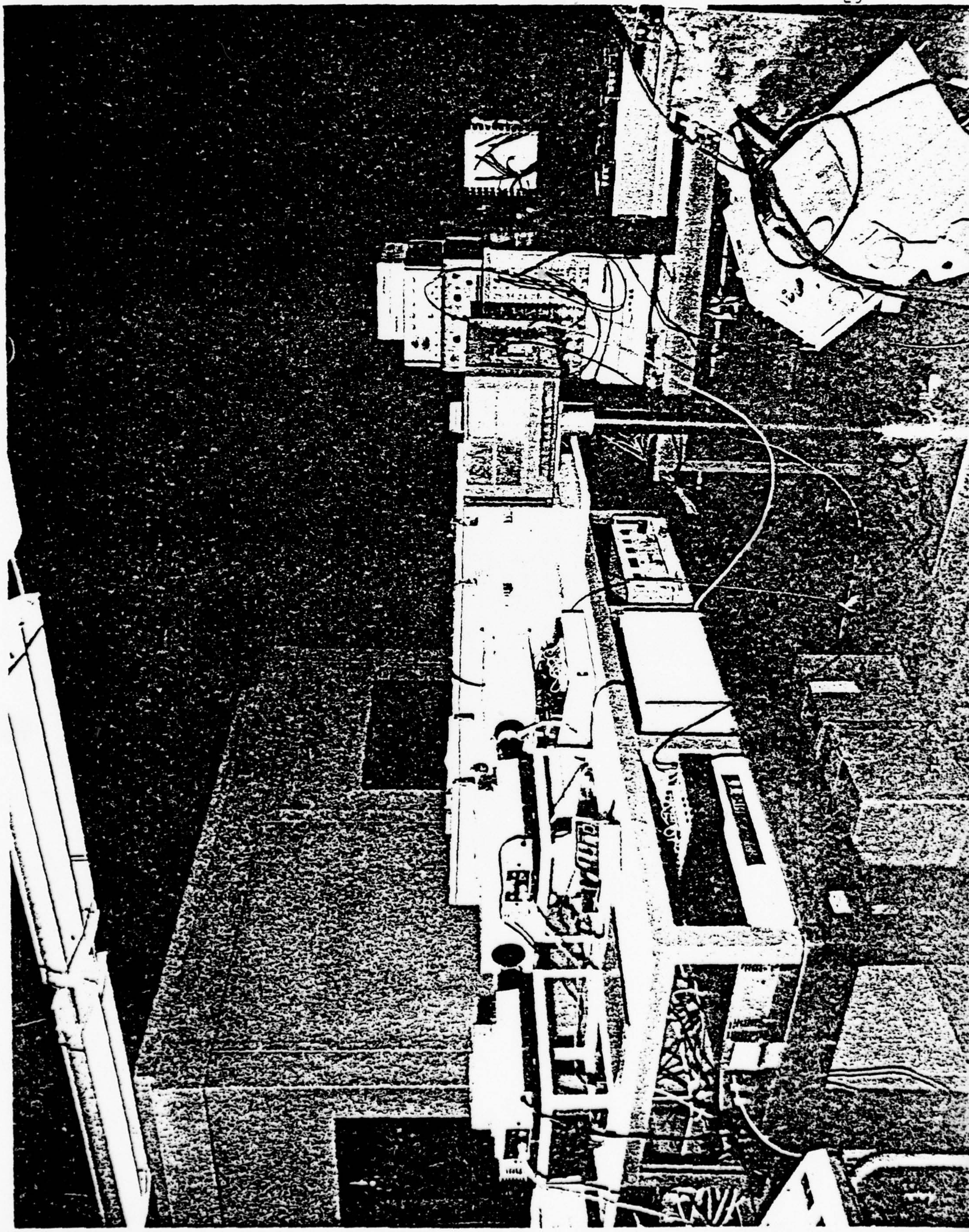
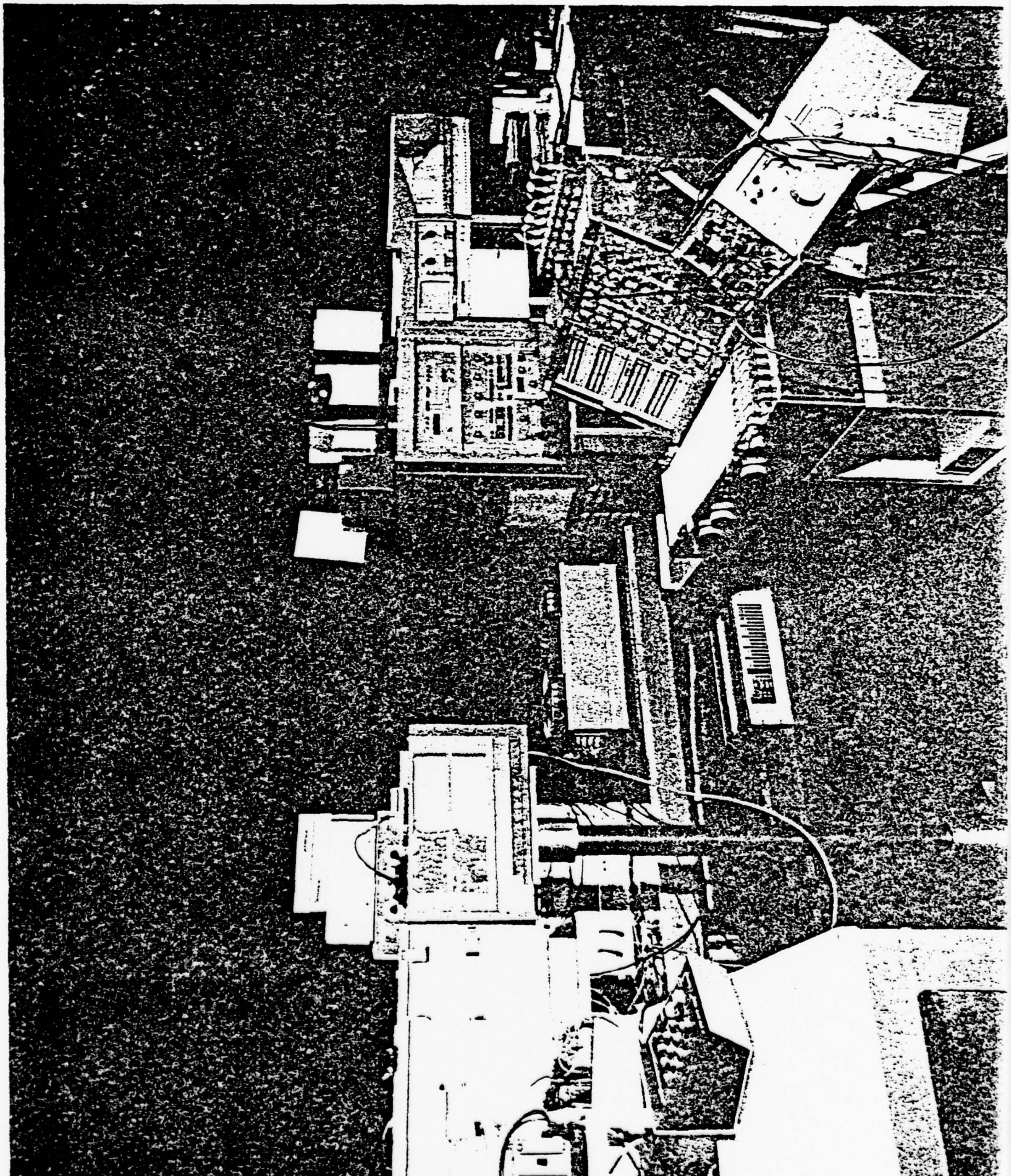


Photo 2. Recording and averaging equipment; polygraph, tape recorder, and averager.



AIR FORCE HUMAN RESOURCES LABORATORY

Research Associates:

Richard J. Bethke, Wright State University

Moira K. LeMay, Montclair State College

Hewitt H. Young, Arizona State University

TEXTURE MODELING AND GENERATION USING PARTIAL DIFFERENCE EQUATIONS

FINAL REPORT

R. J. Bethke

USAF/ASEE Summer Faculty Research Program

August 1978

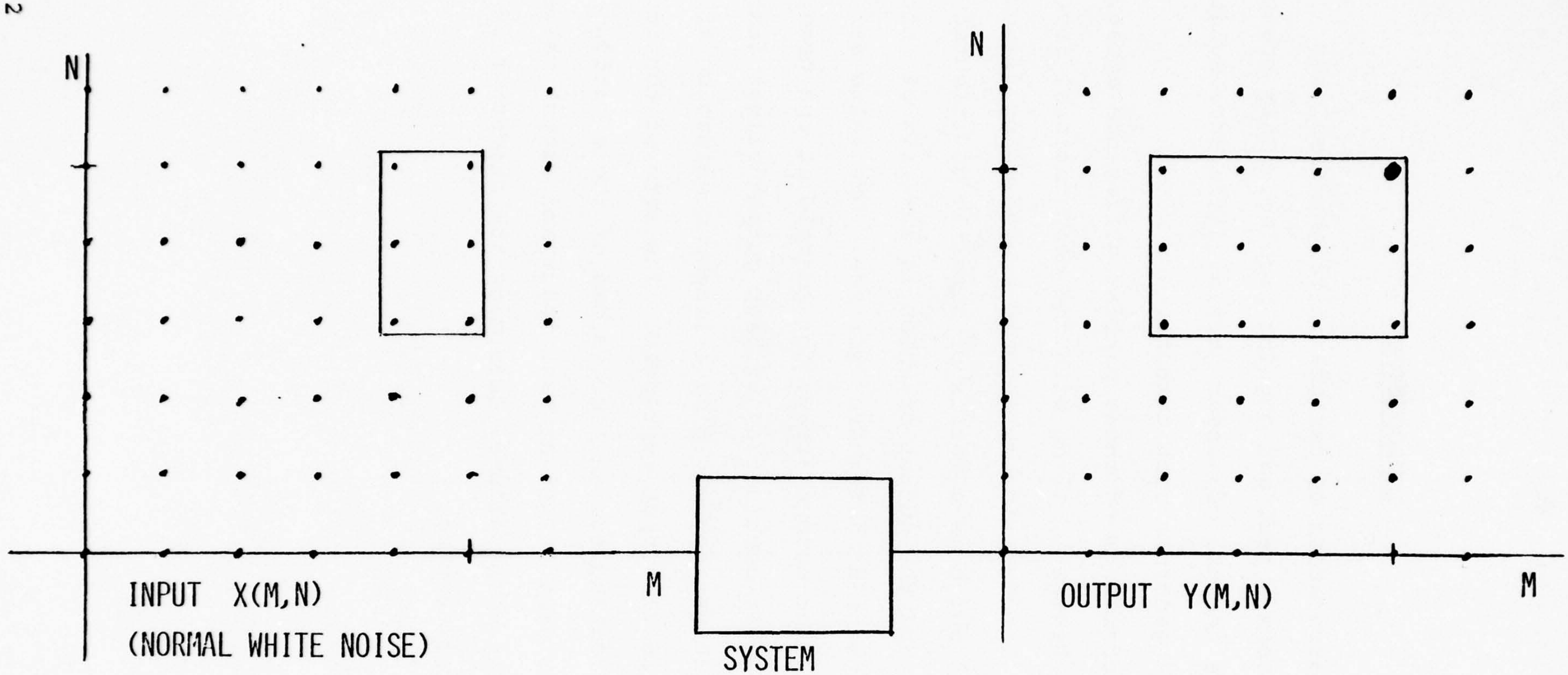
ABSTRACT

The use of linear partial difference equations for generation of texture in flight simulator and other displays is investigated. The ability of these equations to generate a wide variety of synthetic textures is demonstrated, as is the inate flexibility and versatility of this approach to the texturing problem. Generation of perspective views of textured surfaces, using perspective guided equation root trajectories, is demonstrated. The use of partial difference equations in modeling actual textures is explored. The ability of autoregressive forms of these equations to extract large amounts of serial information from these textures is shown.

INTRODUCTION

The incorporation of texture in flight simulator scenes has been a real and ongoing problem. This study investigates the use of linear partial difference equations to model and generate such texture.

The general form of these partial difference equations is shown in Figure 1, along with the computational representation. The equation represents a system, or filter, the input to which is a matrix of normally distributed white noise and the output of which is a matrix of intersities representing a textured surface. The value at any point in the output array, for example at the heavy dot, is a linear combination of past output values (those in the output rectangle), plus a linear combination of the values in the input rectangle. The size of the input and output rectangles, and the values of the equation coefficients used to weight past input and output values in the linear combinations, determine the character of the generated texture.



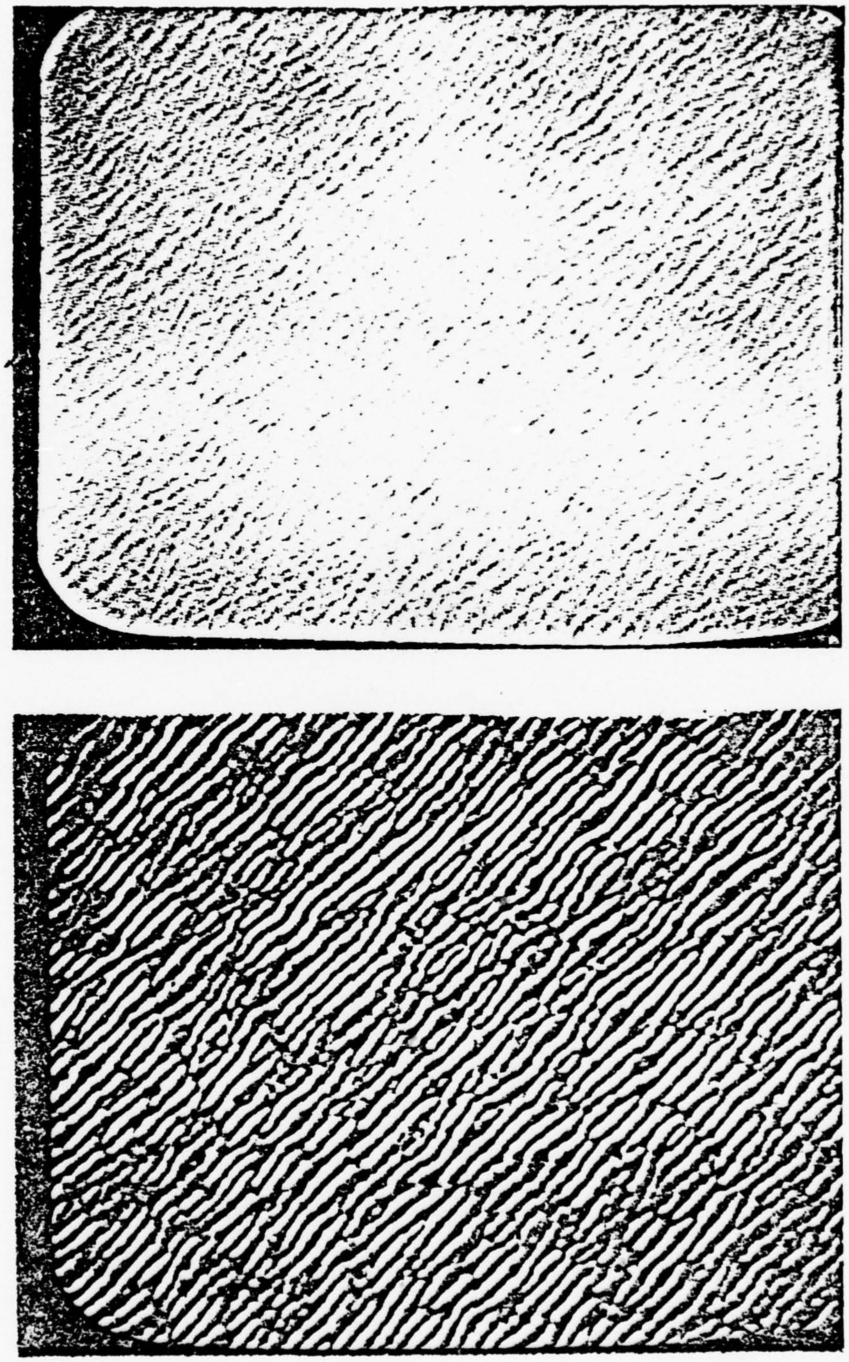
$$\sum_{K=0}^{M_1} \sum_{R=0}^{N_1} A_{KR}\, Y(M - K, N - R) = \sum_{K=0}^{M_2} \sum_{R=0}^{N_2} B_{KR}\, X(M - K, N - R)$$

Figure 1 Computational Configuration

Some Examples of Generated Texture

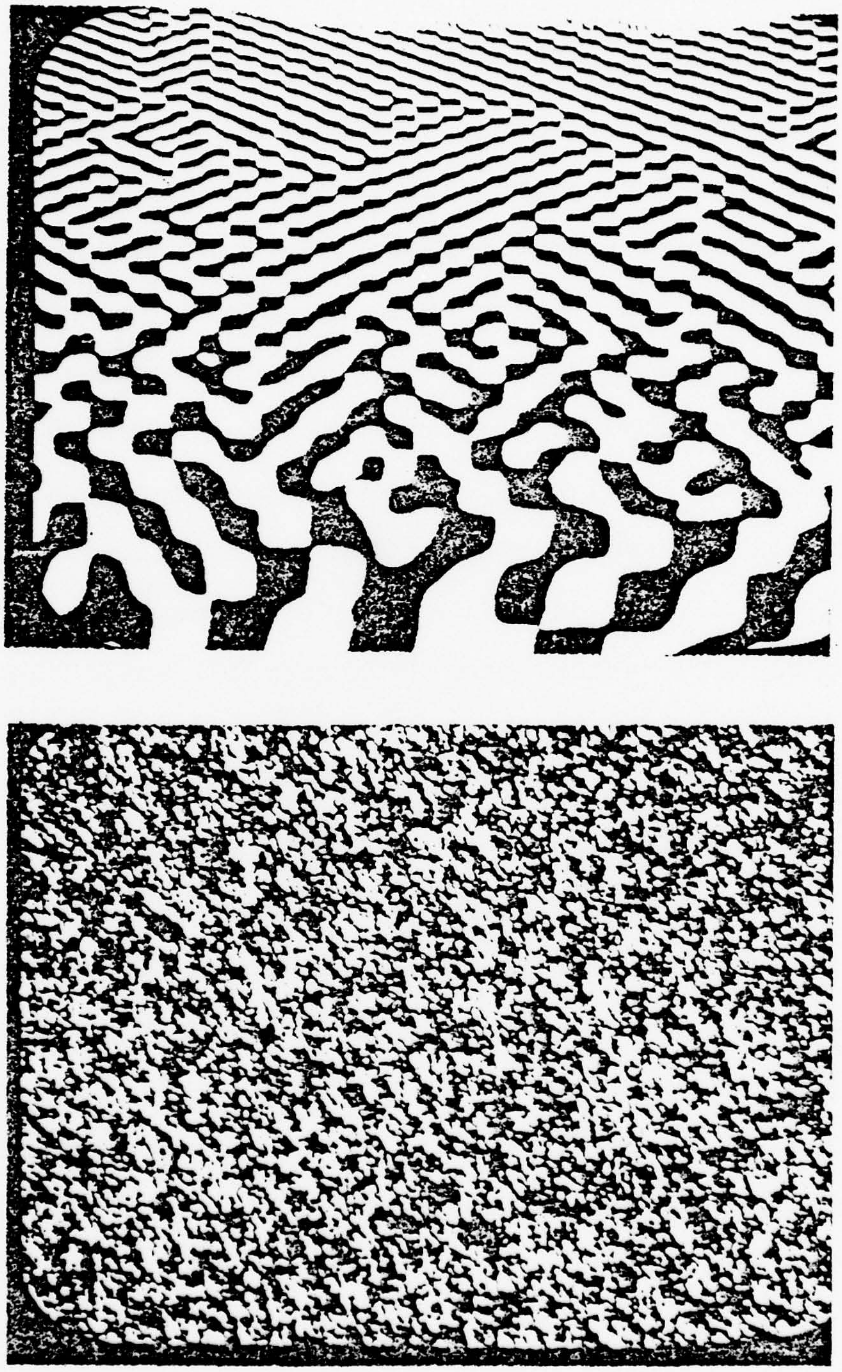
Figures 2 and 3 show some examples of synthetic texture generated by partial difference equations. These examples are included only as a small indication of the wide variety of possible textures. The textures in Figure 2 were produced by equations very similar in form but quite different in amplitude gain. This demonstrates that gain is an important factor in perceived texture. Many people feel the upper picture looks like an aerial view of the sea.

Figure 3 shows a perspective view of a highly over-driven case and the output of a partial difference equation model fitted to a picture of trees. The trees model has had contrast shadowing enhanced by using a D.C. bias to clip the grey levels at the black end of the scale.



4

Figure 2 Examples of Generated Texture



5

Figure 3 Examples of Generated Texture

Generation of Texture in Perspective

Perspective was incorporated in the texture by controlling spatial frequency as a function of perspective. This was done during the actual texture generation as opposed to the more traditional approach of modifying existing texture. It is felt this approach may be more direct and hence more amenable to high speed generation.

The computer program TEXGEN3 was developed to test the feasibility of perspective controlled generation. It utilizes a separable two dimensional filter, $H(Z_1, Z_2) = H(Z_1)\ H(Z_2)$. One filter, $H(Z_1)$, controls spatial frequencies along the raster lines (horizontal), while the second, $H(Z_2)$, controls frequencies perpendicular to the raster lines (vertial). TEXGEN3 introduces perspective due to pitch with constant spatial frequencies along any given raster line.

Apparent spatial frequencies at any given point of the textured surface are proportional to the viewing distance to that point. These frequencies are further altered by surface for shortening due to the surface not being perpendicular to the line of view. For computational expedience TEXGEN3 assumes a perpendicular view along any given raster line and thus the frequencies in any given raster line are controlled by a single viewing distance. Frequencies in the direction perpendicular to the raster lines are controlled both by viewing distance and forshortening.

Control of spatial frequencies is carried out as shown in Figure 4, using a single filter pole for illustration.

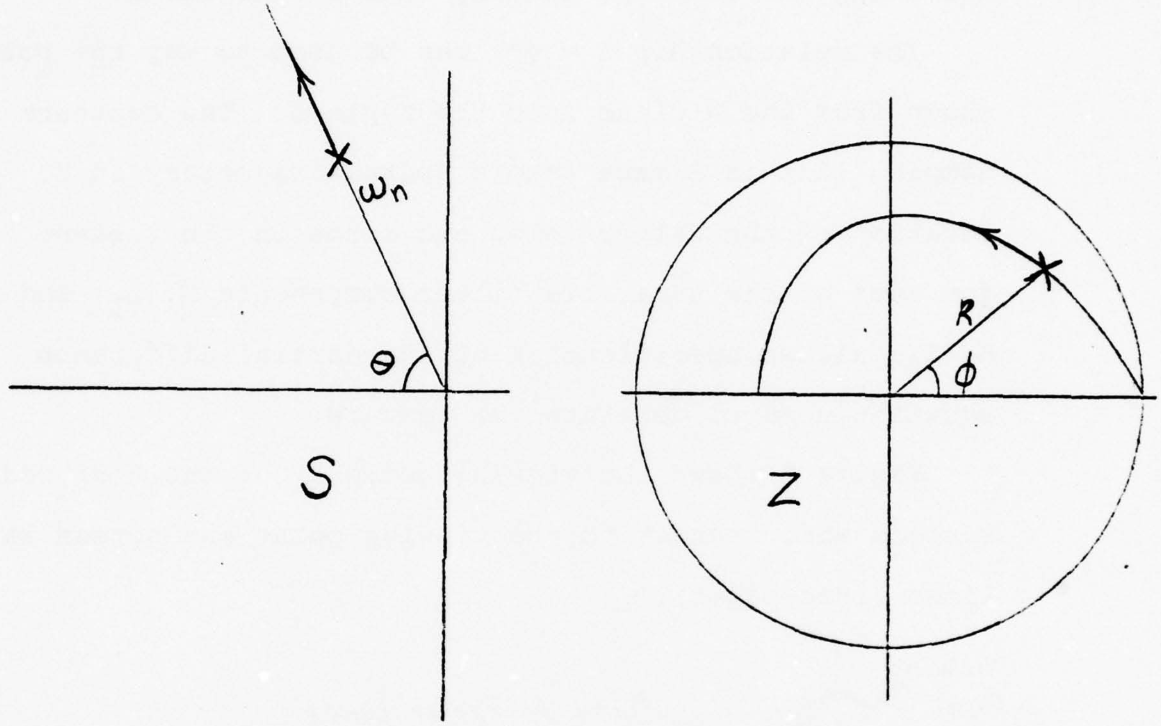


Figure 4 S Plane - Z Plane Mapping

The pole is shown located in the S plane at some natural frequency, W_n. TEXGEN3 uses this reference frequency as the natural frequency along the raster line at the bottom of the viewing screen. In a large filter there are many such reference frequencies in the bottom raster line.

The damping ratio for the pole shown, ζ, determines the angular position, θ, where $\zeta = \cos\theta$. As spatial frequency increases due to perspective, the pole will move out along the constant damping radial line shown.

The relationship $Z = e^{sT}$ can be used to map the pole shown from the S plane into the Z plane. The constant damping line in S maps into a spiral trajectory in Z. Location of the filter poles and zeros in the Z plane for each of the separable filter components $H_1(Z_1)$ and $H_2(Z_2)$ allows specification of the partial difference equation used to generate the texture.

Figure 5 shows the viewing geometry of the textured surface with respect to the viewing point and screen raster lines (into paper).

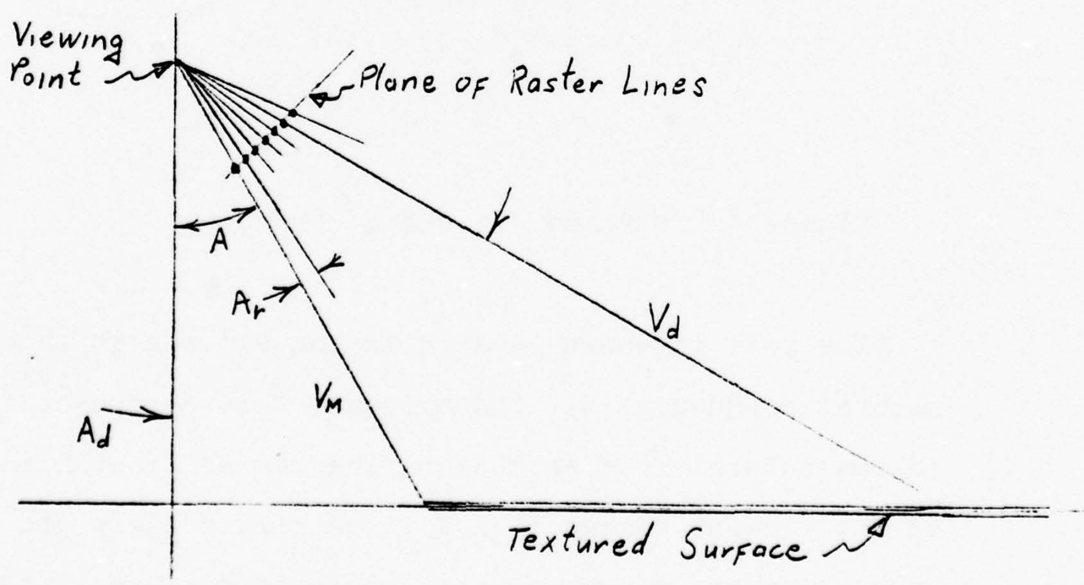


Figure 5 Viewing Geometry

The viewing distance through the bottom raster to the textured surface is Vm at an angle of A radians. The viewing lines to the surface through the raster lines are assumed to be separated by equal angles Ar. The viewing distance through any given raster line to the surface is defined as Vd. Viewing distance ratio, Vr, is defined as Vd/Vm.

The frequency along the raster corresponding to distance Vd and reference natural frequency Wn is a linear function of viewing ratio and is given by

$$Wn' = WnVr$$

This is the radius to the S-plane pole in Figure 4 for the given Vr. This radius and associated damping ratio map into the Z-plane as

$$Z = e^{sT} = e^{(W_n' \cos \theta + W_n' j \sin \theta)\, T}$$

$$= R' \, e^{j\phi' T}$$

where the radius and angle in the Z plane are given by (letting T=1).

$$R' = e^{Wn' \cos \theta} = e^{Wn\, Vr \cos \theta}$$

$$\phi' = Wn' \sin \theta = Wn\, Vr \sin \theta$$

In terms of the original reference Z-plane position

$$R' = R^{Vr}$$

$$\phi' = Vr\phi$$

The frequency perpendicular to the raster lines, W_n'', is effected by not only viewing ratio but by surface forshortening. It can be expressed as

$$Wn'' = Wn'/\sin(\frac{\pi}{2} - Ad) = Wn\ Vr/\sin(\frac{\pi}{2} - Ad)$$

The corresponding Z-plane coordinate is

$$R'' = e^{Wn''} \cos\theta$$

$$\phi'' = W_n'' \sin\theta$$

or

$$R'' = R^{Vr/\sin(\frac{\pi}{2} - Ad)}$$

$$\phi'' = (Vr/\sin(\frac{\pi}{2} - Ad))\phi$$

TEXGEN3 computes Z-plane pole and zero locations based on perspective geometry, computes the corresponding 2-dimensional difference equation, and enumerates texture values along each raster line from a white noise input. Figure 6 shows examples of the resulting texture.

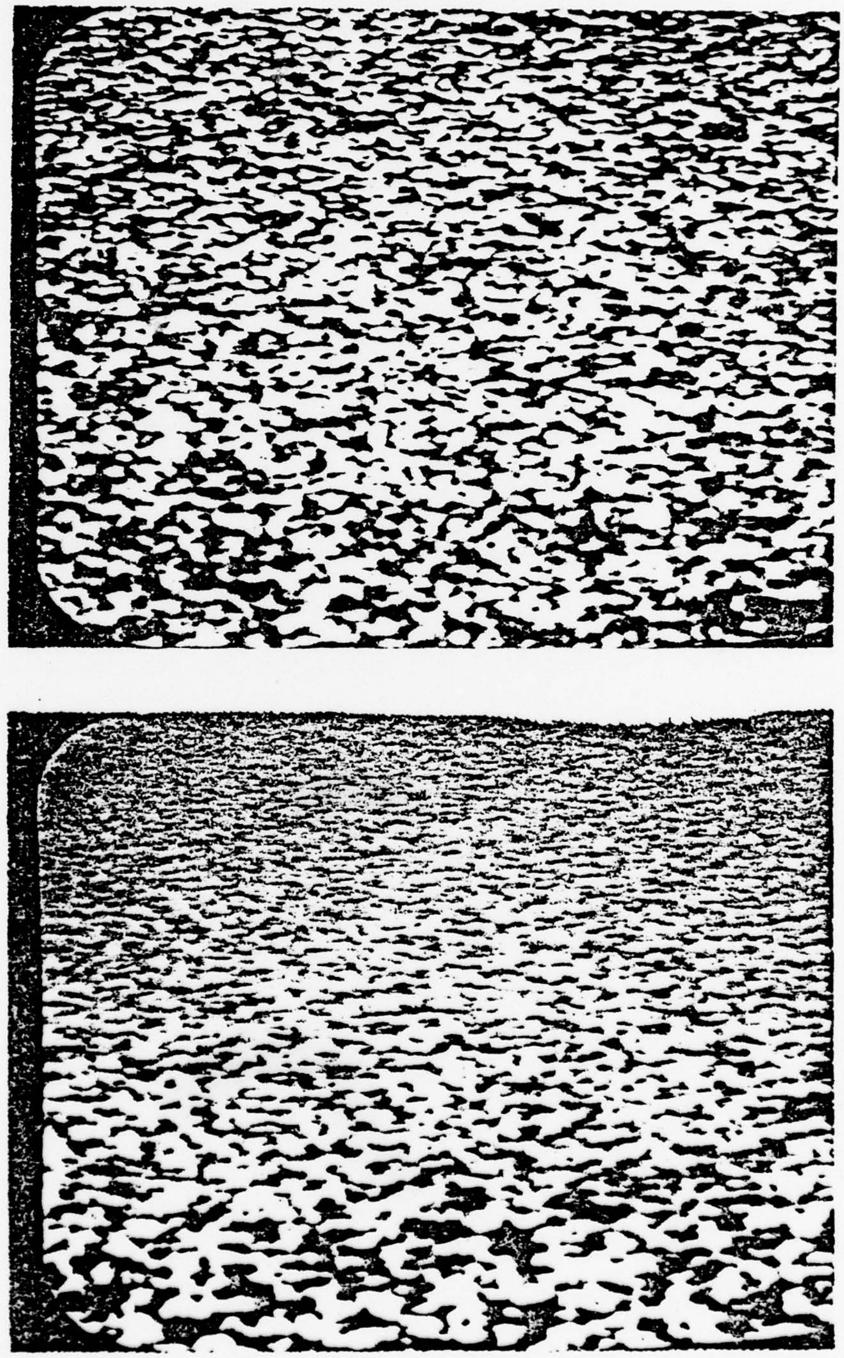


Figure 6 Texture with Perspective

In general the results produced by TEXGEN3 were good. One shortcoming, however, was the lack of control of aliasing. Highly oblique views could push poles and/or zeros along their Z-plane trajectories and across the negative real axis. At this point aliasing occurs and higher frequencies will appear as lower frequencies. There are several remedies for the aliasing problem. Software limits could be imposed on the pole and Zero trajectories preventing their crossing of the negative real axis. Perhaps a better solution is the use of bilinear transformation mapping, $S = Z-1/Z+1$, between the S and Z-planes instead of $Z=e^{ST}$. While this will result in some frequency distortion it is probably the superior solution to the aliasing problem.

Texture Modeling

Modeling of digitized photographs of actual texture was conducted primarily using two scenes stored on TSC tape 00601; trees (File 2), and grass (File 4). The modeling process involves basically three steps; identification, estimation, and diagnostic checking.

In the identification step the data is examined to determine statistical distribution and the form and size of the model which might be used to represent it. Correlation functions are useful in this step.

The estimation step involves finding the least squares estimates of the difference equation coefficients using linear or nonlinear regression. Diagnostic checking of the fitted model is then done through examination of the residuals (errors) for mean, variance, statistical distribution, and correlation.

The computer program LINREG was written to conduct the above modeling steps. This program computes mean, variance, histogram and auto correlation of the input data, computes the model coefficients via linear regression, and computes the residuals their mean variance, histogram and autocorrelation.

Figure 7 shows the data to be modeled, trees and grass, top and bottom respectively. Figures 8 and 9 show the

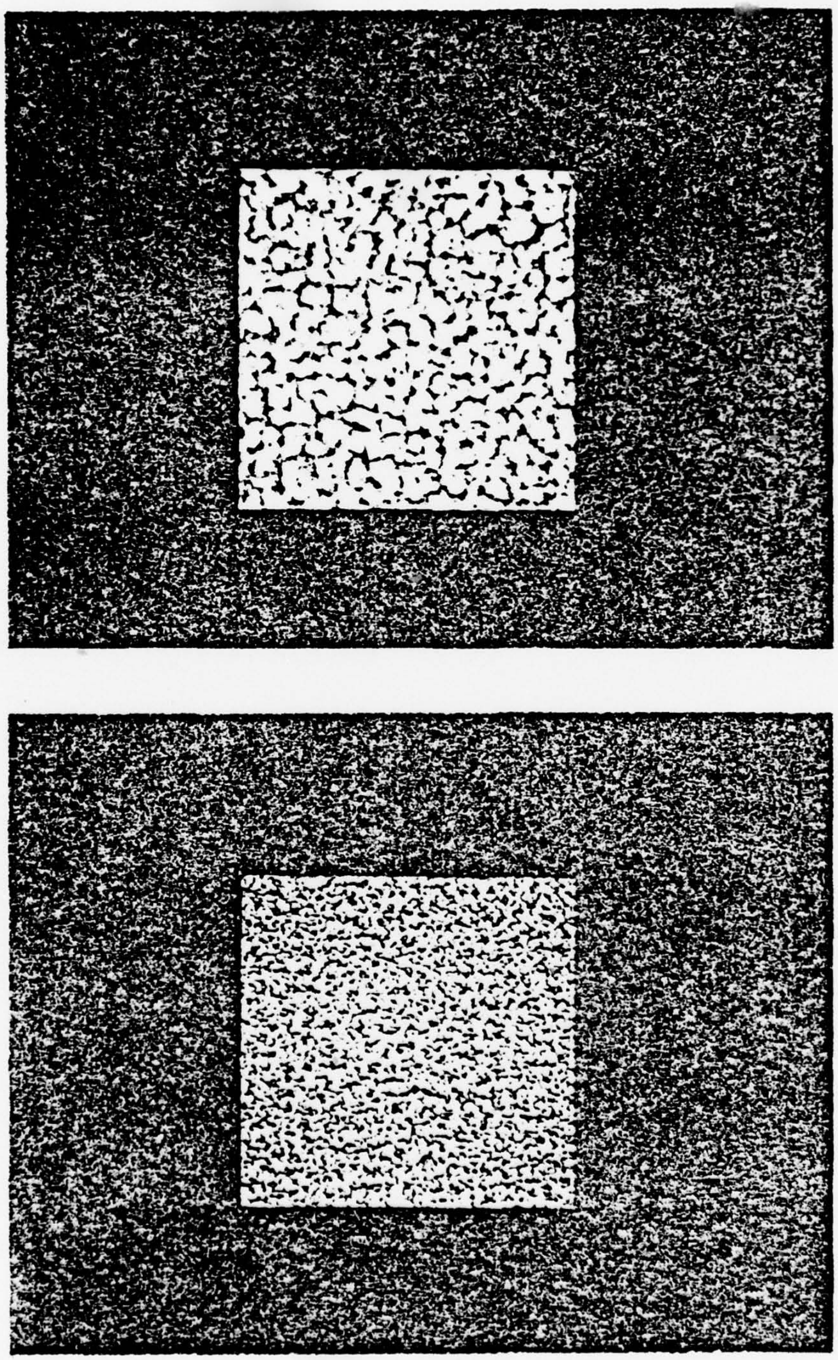
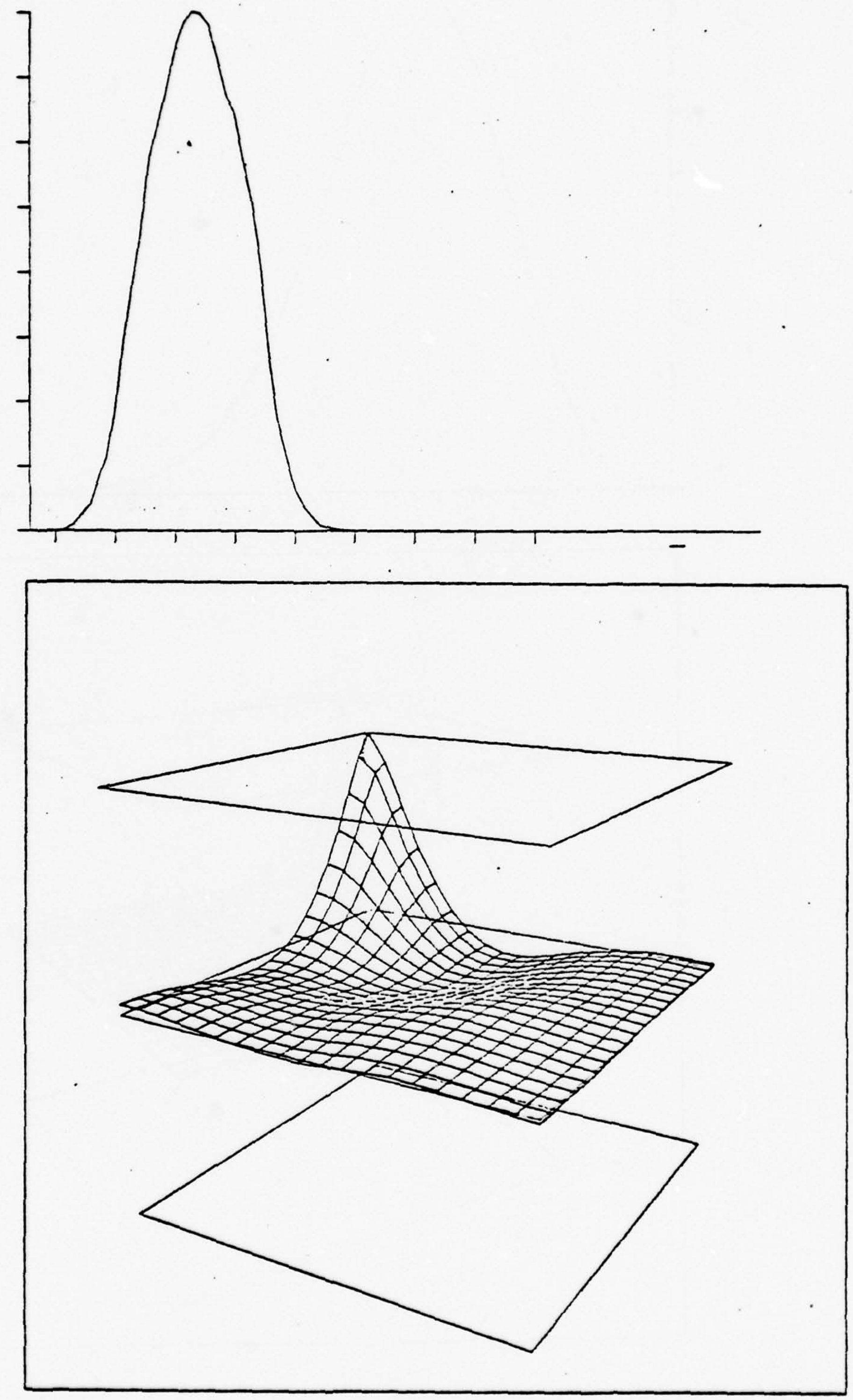
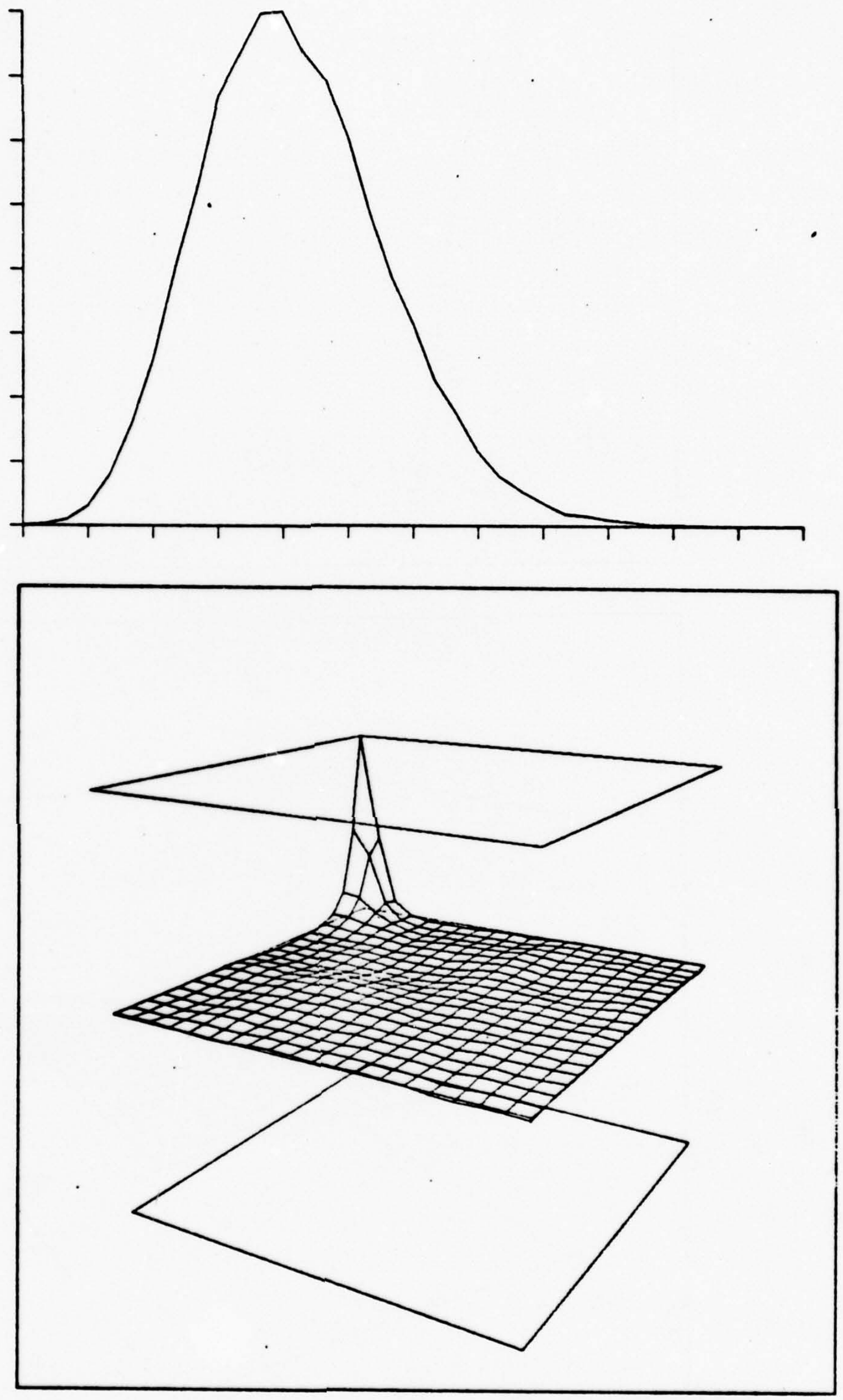


Figure 7 Data - Trees (top); Grass (Bottom)



15

Figure 8 Histogram and Autocorrelation Function - Trees



16

Figure 9   Histogram and Autocorrelation Function - Grass

histograms and autocorrelation functions of this data. The histograms are unimodal and reasonably symetric. The assumption of normally distributed data appears plausible. The autocorrelation functions roll off smoothly without steps or discontinuities. This is a indication that auto-regressive models may be appropriate. That is, models of the form.

$$Y(M,N)=X(M,N)+\sum_{K=0}^{M_1}\sum_{R=0}^{N_1} A_{kr}\, Y(M-K,\ N-R)$$

where the Y's are the picture values, X is the current input (or error), and the A's are the parameters to be estimated ($A_{00}=0$). In this study $M_1=N_1$ for convenience, and the A parameters form a matrix of size MXM.

After estimation of the parameters the X values can be computed and viewed as residuals (errors). The smaller the variance of these residuals the better the fit of the model. Table 1 gives the residual variance and the ratio of this variance to that of the original data. Note that for both grass and trees the residual variance drops rapidly and then slowly as the model size increases. It appears that little is gained after M is three or four. Note also the large difference in the residual variance to data variance ratios, between trees and grass. These ratios are controlled by both model fit and the amount of information in the data. When the model fits and extracts the serial information from the data, the size of the ratio indicates how uncorrelated or noisey the original data was.

	TREES		GRASS	
Model Size M	Residual Variance	Ratio	Residual Variance	Ratio
1	27.750	1	18.082	1
2	3.365	.1212	11.531	.6377
3	2.673	.0963	11.070	.6122
4	2.643	.0952	11.032	.6100
6			11.009	.6088
7	2.624	.0945	11.007	.6087

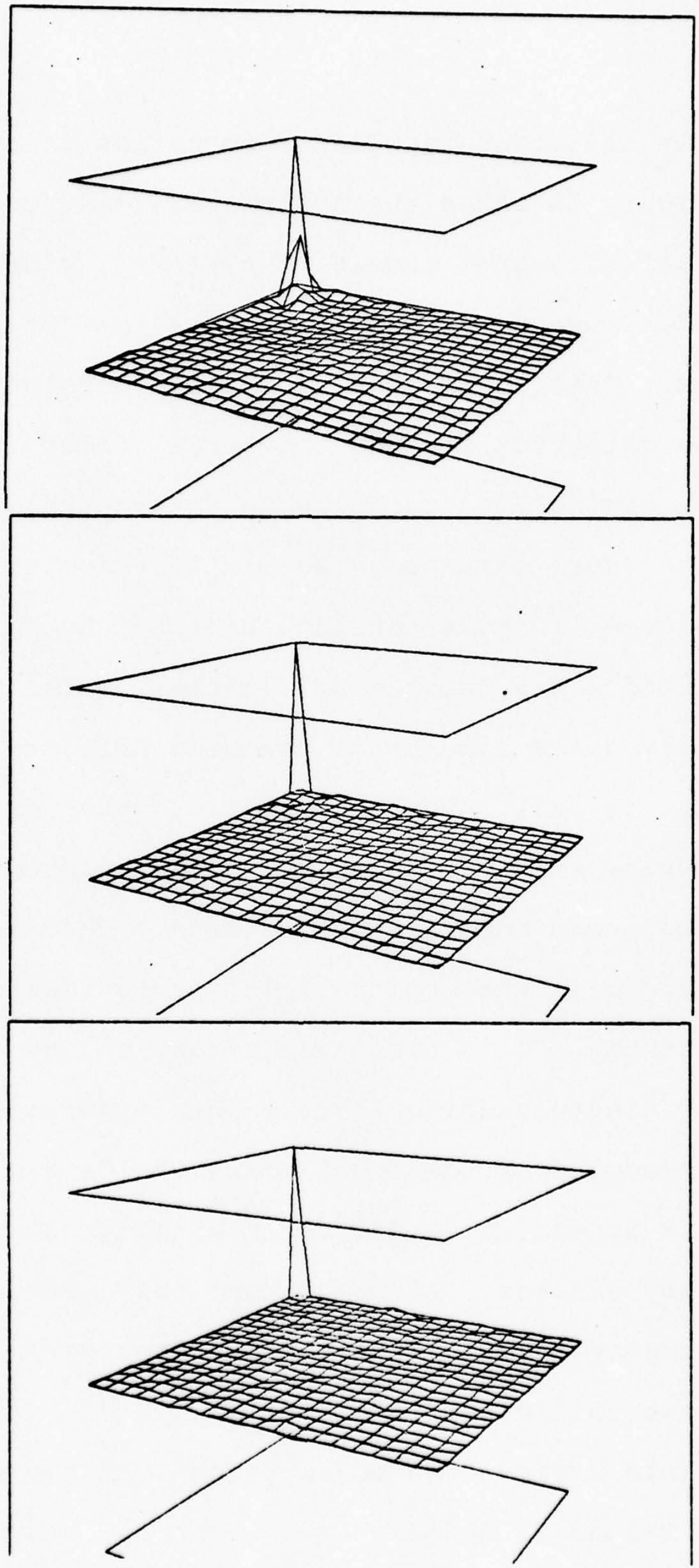
Table I Residual Variances and Ratios

Clearly the grass has less information in it than the trees.

Figure 10 shows the autocorrelation functions of the residuals for tree models of size M=2, 4 and 7. Note the relative absence of serial correlation for the two larger models. This shows reasonably good model fit, the model having extracted most of the serial information from the data, leaving the residuals quite uncorrelated. The models for the grass data produced similar results.

Figure 11 shows texture generated by a M=4 tree model (top) and a M=6 grass model (bottom). The grass texture is visually quite similar to the original data, the tree texture is not. Although the tree model extracted most of the serial information in the data, it cannot account for the nonlinear behavior in the data. That is, the light and dark areas in the original picture as seen to behave differently. This nonlinear effect is, regretably, a highly distinguishing feature when viewing the tree data. Its absence from the model makes the data and the generated texture appear quite different although statistically they are very similar. Although nonlinear models are outside the present study, it is thought they could successfully increase this visual similarity.

Table 2 gives the model parameters for a M=7 tree model and a M-6 grass model.



M=2

M=4

M=7

Figure 10 Autocorrelation Functions of Residuals

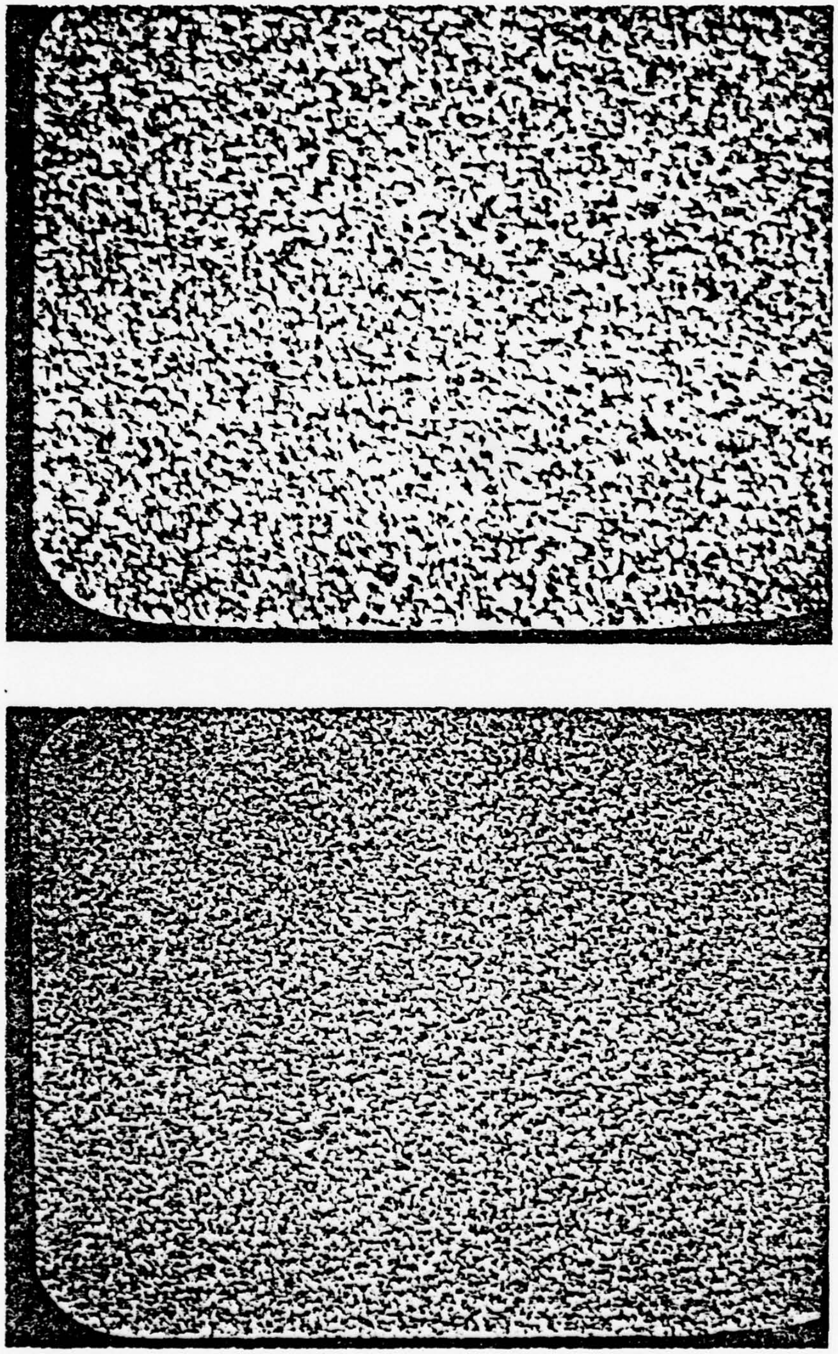


Figure 11 Texture Generated From Models

M=7 Trees

.00000000	.61113733	-.11879611	.03044363	-.01586614	.00206642	-.00905756
.86565679	.01666488	-.17653787	.03675846	-.01462927	.00758305	-.00939088
-.23524392	-.21999121	.00178588	.04378932	-.01231859	.01148294	.00682854
.02921797	.02322938	.08741724	-.00524750	-.01214743	-.00766214	-.01081096
.00142878	.03114546	-.01963389	-.03773968	.01717842	.01057243	-.00061017
.00420561	-.01633501	-.01660761	.01363510	.00271225	-.01384480	-.00081348
-.01232954	-.00013540	.00392442	.00567270	-.00736332	.00170481	.00173855

M=6 Grass

.00000000	.39518595	-.12622171	.03879564	-.01553615	.00223015
.46471440	.05769032	-.05556772	.02263266	-.01872677	.00209600
-.11855920	-.07300908	.01177477	.01137510	-.00851733	-.00124906
.03261628	.03511664	-.00243367	-.00052947	-.00025338	-.00372626
-.00809476	-.01234464	-.00676340	.00333022	.00049317	-.00313751
-.00003920	-.00897798	-.00053842	-.00650805	.00066987	-.01360138

Table 2 Model Parameters

CONCLUSION

This study has demonstrated the ability of two dimensional partial difference equations to generate a wide variety of textures which can be used in flight simulator and other displays. Perspective views of textured surfaces can be created by controlling the difference equation parameters as the surface is being generated. These results may have important bearing on high speed display applications.

The modeling portion of the study showed that low order autoregressive models were capable of extracting most of the serial information from the texture data investigated. Lack of fit seemed to be caused more by nonlinearities in the data than by departures from normality or the lack of moving average terms in the model.

COMPLEXITY JUDGEMENTS OF COMPUTER GENERATED SIMULATIONS OF ELECTRO-OPTICAL DISPLAYS

Moira LeMay
Montclair State College
Upper Montclair, New Jersey

August 1978

AIR FORCE HUMAN RESOURCES LABORATORY
ADVANCED SYSTEMS DIVISION
WRIGHT-PATTERSON AIR FORCE BASE, OHIO 45433

Submitted in fulfillment of
The 1978 USAF-ASEE Summer Faculty Program

Sponsored by

THE AIR FORCE OFFICE OF SCIENTIFIC RESEARCH

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

COMPLEXITY JUDGEMENTS OF COMPUTER GENERATED SIMULATIONS

OF ELECTRO-OPTICAL DISPLAYS

Moira LeMay
Montclair State College
Upper Montclair, New Jersey

ABSTRACT

In the simulation of electro-optical sensor displays, the extent of realism needed for optimal transfer of training is assumed to be dependent on scene complexity. Therefore, the effect of a number of parameters amenable to change in computer image generation on judged complexity was investigated. A single scene was generated on the computer consisting only of cultural objects and was varied in terms of external edges (shapes of objects), internal edges (doors and windows), noise in the display (study I) and display contrast, the addition of shading and of a Gaussian transfer function (study II). Complexity ratings of the scene were obtained from three groups of subjects, each under slightly different conditions.

The results showed that noise was an extremely important, though somewhat extraneous, determinant of judged complexity, and its effects could be only partially controlled by the introduction of an instructional set.

Once a scene had been modeled on the computer and a barely sufficient number of edges had been used to outline the objects, further introduction of edges to object outlines made no significant difference. However, the addition of internal edges in the form of doors and windows significantly increased complexity ratings. Contrast, shading, and the transfer function had no effect on perceived complexity.

COMPLEXITY JUDGEMENTS OF COMPUTER GENERATED SIMULATIONS OF ELECTRO-OPTICAL DISPLAYS

Moira LeMay

Montclair State College

A perennial problem in the use of simulators has been the degree of realism necessary to produce optimum transfer of training. In the simulation of electro-optical (E/O) sensor displays, such as Low Light Level Television (LLLTV) and Forward Looking Infrared (FLIR), the extent of realism has been assumed to be dependent on scene complexity. The definition of complexity, however, is by no means clear and easy. In studies of E/O display simulation, it is often assumed to be related to the level of detail which can be displayed, or the number of edges, contours, transfer function definition, and other variables which can be added to the display at increasing cost (Bunker and Heeschen, 1975; Bunker and Heartz, 1976; Bunker and Ferris, 1977). These factors certainly increase the complexity of producing a display; their relationship to the complexity of a display as perceived by an observer is the subject of the present investigation.

Perceived complexity, as it relates to various form characteristics, has perhaps been studied most commonly using a seven-point rating scale. Attneave (1957), in an early study using random polygons generated according to the technique suggested by Attneave and Arnoult (1956), found that judgements on such a scale were related mainly to the number of sides or turns in a polygon, with symmetry playing a lesser role. Vanderplas and Garvin (1959), in turn, used Attneave's results to construct figures of varying complexity, and found that increasing complexity was

inversely related to a number of associative responses and correct recognition, and positively related to response time for recognition.

Glanzer and Clark (1964) present a "verbal loop" hypothesis, and relate complexity to length of verbalization used in describing the stimulus, although a cross-cultural repetition failed to confirm this finding (Tone and Muraoka, 1975). In addition, the complexity scale used by Glanzer and Clark, which was experimenter-defined, was found to have no significant relationship to subject-validated complexity ratings, indicating that the determinants of subjective complexity cannot be assumed for a particular set of forms, but must be investigated if the forms are to be used in further work in which complexity is important.

LeMay (1971) and LeMay and Locher (1978) also found that judged complexity was related to number of turns or sides, not only in random polygons, but in lissajous figures and in both outline drawings and photographs of single buildings (houses). In addition, the surface details of a building in the form of doors, windows, and visible structural elements (as in Tudor houses) added to judged building complexity.

The seven-point rating scale has also been used in situations where stimuli are so complex that any other metric would be difficult to apply. Leckart and Bakan (1965) applied it to a series of slides of natural stimuli, including landscapes, single objects, and arrays of objects. They found that higher ratings were related to more exploratory behavior in the same way that Berlyne (1960) had found experimentater-defined complexity of laboratory-constructed stimuli was related to free looking time. LeMay and Aronow (1977) also used the rating technique with scenes from the interiors of buildings and related complexity both to free looking time and to lingering behavior. Saklofske (1975) asked subjects to rate

the complexity of paintings on a seven-point scale, and related complexity to free looking time, and also to the number of questions formulated about the paintings. Dewer and Ells (1977) used complexity ratings on a seven-point scale as part of a semantic differential study of traffic signs, and subsumed it under an "understandability" factor.

De Groot (1978) studied complexity ratings of photographs of LLLTV and FLIR displays, and constructed a scale consisting of six levels of scene complexity for LLLTV and four for FLIR. A content analysis of subject debriefings in this study revealed that the factors that influenced complexity judgements were the number of man-made or cultural objects in the scene, their discriminability, area, detail of structure and sub-structures, sharpness of edges, and the number and directions of lines and edges, although there were some differences between LLLTV and FLIR. Subjects were able to match the sensor displays with photographs of the same target areas with very few errors, most of which occurred at the low end of the complexity scale. High correlations were found between the scalings of the two sensor displays. Because the stimulus materials were unavailable, a further analysis could not be performed. The materials were available for the present study, however, and the analysis is reported later in this paper.

Other scaling techniques have also been used to rate the complexity of various types of stimuli. Chipman (1972) used a method of magnitude estimation to obtain a complexity score for each of 30 checkerboard patterns, and found that perceived complexity was affected by stimulus organization, more organized stimuli being seen as simpler. Chipman and Mendelson (1975) extended this finding to children of various ages, using a method of paired comparisons appropriate for younger age groups.

Perhaps the most comprehensive studies of exploratory and aesthetic behavior have been carried out by Berlyne and his associates (Berlyne, 1958, 1960, 1974; Berlyne and Parham, 1968) with experimenter-defined stimulus complexity frequently figuring as an important independent variable. A positive relationship between stimulus complexity and exploratory behavior has been found so consistently that an increase in such behavior may be taken as indicative of an increase in subjective complexity (Wilson, 1976). However, the parameters that lead to the judgement that a stimulus is complex have not been investigated directly in these studies, and complexity is usually defined in the explicit descriptions of simple and complex stimuli.

Several interesting techniques have been used in an attempt to elaborate a definition of complexity independent of subjective ratings. Pasnak (1969) and Seman, Pasnak, and Tyer (1976) used a technique suggested by Attneave (1954) to quantify the complexity of natural stimuli (human faces and human hands) and Nygard, et al (1964) measured complexity with a stimulus complexity analyzer which scanned object size, object brightness, and object count.

It might be concluded from these studies that subjective complexity is chiefly determined by numbers, either of objects or parts of objects (such as sides), and secondarily influenced by organization. In addition, complexity has effects on other kinds of behavior, particularly exploratory or attentional behavior. One would expect that, if more attention is given to more complex stimuli, they could be recognized and identified more easily than simpler stimuli, although it might be more difficult to detect them. The effect of complexity on detection, recognition, and identification is, of course, the area of greatest interest with regard to the simulation of E/O sensors.

Relatively few studies have investigated the effect of pattern complexity on detection directly, although many have studied detection as a function of similar stimulus variables, or the complexity of the task (e.g., Curran, 1975; Curry, Day, Durmeuger, and Senders, 1975; Singer and Lappin, 1976; Sternberg and Banks, 1970). Uttal and Tucker (1977), however, found a decline in detection performance with increasing complexity. They also found complexity to be a powerful determinant of susceptibility to masking by noise.

The results of studies on the effects of complexity on recognition have been somewhat equivocal. Goedert and Rodwan (1977) and Salmon (1977) found no effect, while Nygard, et al (1964) found a relationship between some of the measures obtained from their stimulus complexity analyzer for target backgrounds and target recognition. Vanderplas and Garvin (1959a) found that increasing complexity led to a decrease in correct recognitions and an increase in response time. Meyer (1973) however, in a somewhat different study involving architectural rather than laboratory-generated stimuli, found that subjects recognized the function of a complex building more easily than that of a simpler one.

There is some theoretical justification, and some evidence, for the notion that complex stimuli may be more easily identified than simple stimuli. Gubson, Gibson, Pick and Osser (1962) proposed that a distinctive part of a form (a "distinctive feature") is the discriminative stimulus for that form, and found that children learn these distinctive features earlier than they learn other parts of a form. Homa and Coury (1977) found that identification of a distinctive feature may be enhanced, although only at the expense of other parts of a stimulus array. If distinctive features are the basis of identification, then adding complexity should aid identification by adding more such features, and the selective

attention reported by Homa and Coury should enhance the effect. Indeed, Pasnak (1969) working on Gibson's hypothesis, found that increasing complexity does, in fact, increase the probability of correct identification, along with some confirming evidence for the distinctive feature notion (Pasnak, 1971; Seman, Pasnak, and Tyer, 1976)

In view of the above findings, it will probably be useful in the future to study more comprehensively the effect of stimulus complexity on detection, recognition, and identification, since the required degree of complexity may depend on the task to be performed. With this in mind, the present study was designed to define some physical parameters of judged scene complexity which can be modeled using computer generated images so that, in future task studies, the most relevant parameters may be varied.

Since the literature review had indicated that the most important factors in determining judged complexity in a display were the number of man-made objects in the scene, and the number of edges, both external (sides of a building) and internal (surface detail of a building) it was decided to generate on the computer a single scene consisting only of man-made, or cultural objects and to vary that scene in terms of external and internal edges. It was also surmised that noise in the display might affect judgements by masking some of the edge detail, as had been found by Uttal and Tucker (1977) so noise was introduced as a third variable. It was hypothesized that the addition of both internal and external edges would lead to judgements of greater complexity and that noise, because it masked some of the detail, would interact with edges.

There are also several other variables, about which the literature says virtually nothing, which can be added to a computer generated display

and might affect complexity judgements. There are: an edge transfer function, which softens the edges to simulate E/O displays; the amount of contrast in the display, which should have the same effect as noise; and the introduction of shading, which simulates a curved surface on appropriate objects. These were studied in a second series of variations on the original scene. It was somewhat difficult to predict the effect of the transfer function and shading although, in the sense that the former may mask some detail and the latter add some detail, they may affect complexity judgements accordingly.

Supplementary Analysis of Complexity Ratings of LLLTV and FLIR Photographs

The study by de Groot (1978) mentioned earlier reported on complexity ratings for a set of 16 photographs of LLLTV and FLIR scenes in the WPAFB area. The parameter most frequently mentioned in a content analysis of debriefing comments was number of man-made, or cultural, objects in the scene. Also mentioned was the proportion of the scene occupied by cultural objects and the detail of outline shape and building surfaces. The latter two variables are the subject of the present report. The number and area of man-made objects was considered a characteristic of the scene itself rather than of the method by which it is viewed or simulated and, as such, is of somewhat limited interest. However, since it is probably the most important single determinant of scene complexity, a further analysis was undertaken.

De Groot suggests a number of scene parameters which might be quantified, such as the number of separate single entities, ignoring sub-structural

components; the number of separate entities including sub-structural components; the ratio of the area of man-made objects to the area of foliage/terrain and sky; and the number of straight and curved lines which make up the visible portion of man-made objects. The actual quantification of lines and entities, however, turns out to present difficulties in terms of reliability since the outlines of buildings sometimes make it impossible to tell whether there are one or two structures, and it may be difficult to define a cultural object in ways that will be meaningful for complexity judgements. For example, is each landing strip of a runway a single object or is the entire inter-connected set of landing strips of a runway one object? The same question may be used concerning a row of parked cars, a divided highway, an apartment complex, etc. Nevertheless, an attempt was made to count the number of objects in each of the LLLTV and FLIR scenes in a reliable fashion and correlate this with the complexity rating obtained by de Groot for each scene. This data is presented in Table 1 and the correlation between number of objects and complexity ratings for the LLLTV photos was .74; for the FLIR photos, r = .69. Both of these are significant (P < .01) indicating that the number of cultural objects is, in fact, an important determinant of complexity judgement.

Since the object count was felt to be somewhat unreliable, a measure was taken of the area occupied by man-made structures in each of the scenes. To avoid some of the ambiguity in the LLLTV and FLIR photos, measures were taken from daylight photographs of the same scenes. A clear plastic overlay ruled in square millimeters occupied by man-made structures was simply counted. This resulted in the area measure reported in Table 1.

TABLE 1

Numbers and Areas of Cultural Objects and Complexity Ratings

for LLLTV and FLIR Photographs of 16 Scenes

Scene Number	Number of Objects (LLLTV)	Complexity Rating (LLLTV)	Area Occupied By Cultural Objects	Number of Objects (FLIR)	Complexity Rating (FLIR)
1	12	6.6	4371	10	6.1
2	8	5.6	1722	25	5.6
3	7	5.6	2480	17	5.5
4	12	5.6	2613	20	6.0
5	8	6.9	4962	11	6.1
6	3	2.3	1059	10	2.9
7	0	1.1	204	5	1.0
8	3	1.9	407	4	1.6
9	5	3.4	3466	12	3.4
10	10	4.4	3275	7	3.0
11	7	3.4	4284	15	4.1
12	2	2.2	1117	14	2.5
13	4	3.8	638	19	5.1
14	7	4.1	1980	26	5.9
15	2	3.0	910	5	3.8
16	3	1.7	836	8	2.9

The correlation between the area occupied by cultural objects and complexity ratings was .73 for LLLTV photos and .56 for FLIR photos ($P < .01$) confirming the above result. The percentage of total scene areas occupied by terrain, as opposed to sky area, was also calculated and correlated with the amount of cultural area but this lowered the coefficients indicating that subjects "lumped" or generalized the sky area in with foliage/terrain as background for the cultural objects.

These results simply serve to confirm the earlier findings of de Groot (1978) that the number of cultural objects is the main determinant of scene complexity, relatively independent of the natural environment, which is seen as ground against which cultural objects stand out as figure. If it had been possible to count numbers of edges and surface detail of buildings, even more of the variance in complexity ratings might have been accounted for. Thus, it follows that the next step is to remove the background and study the detail of building shapes and surfaces. This is the aim of the study reported below.

METHOD

Subjects

Subjects were volunteers from several classes in psychology and business at Wright State University. They were tested in three groups. The first group consisted of 21 students (14 female and 13 male) tested in small groups ranging in size from one to eight subjects. Group 2 consisted of 16 subjects (11 male and 5 female) tested all at once in a bright, daylit room. Group 3 consisted of 10 subjects (8 male and 2 female) tested under the same circumstances as Group 1.

Stimuli

The stimuli were two sets of slides taken from the computer image generation model described by Bunker and Heeschen (1975). The same basic scene was simulated in all slides. It depicted an aircraft landing strip with three planes on it, a row of four hangars, and several round storage tanks, as well as several other buildings. It was based on an actual scene at WPAFB but was varied to fit experimental requirements. No landscape or natural detail was included. These stimuli are shown in Figures 1 through 9.

The first set of slides varied in three parameters: number of external edges, number of internal edges and amount of noise. There were three levels of each parameter yielding a 3 x 3 x 3 factorial design with 27 stimuli. The number of external edges may be considered equivalent to the number of sides in a random polygon. This was varied by representing buildings at the lowest level by a simple rectangular box, viewed in perspective, and adding external edges so that, for example, at the second level, the hangars had a standard peaked roof (2 edges) and at the third level, the hangar roof appeared curved (5 edges). The number of internal edges meant the addition of doors and windows represented by blocks at the second level and with some internal detail such as window panes at the third level. A random number generation algorithm computes numbers for noise simulation, which can be applied to the scene with any desired strength. Noise levels used in the present study were 0, 25% and 50%.

The second set of slides varied in four parameters: number of external edges, as before, amount of contrast in the image and the

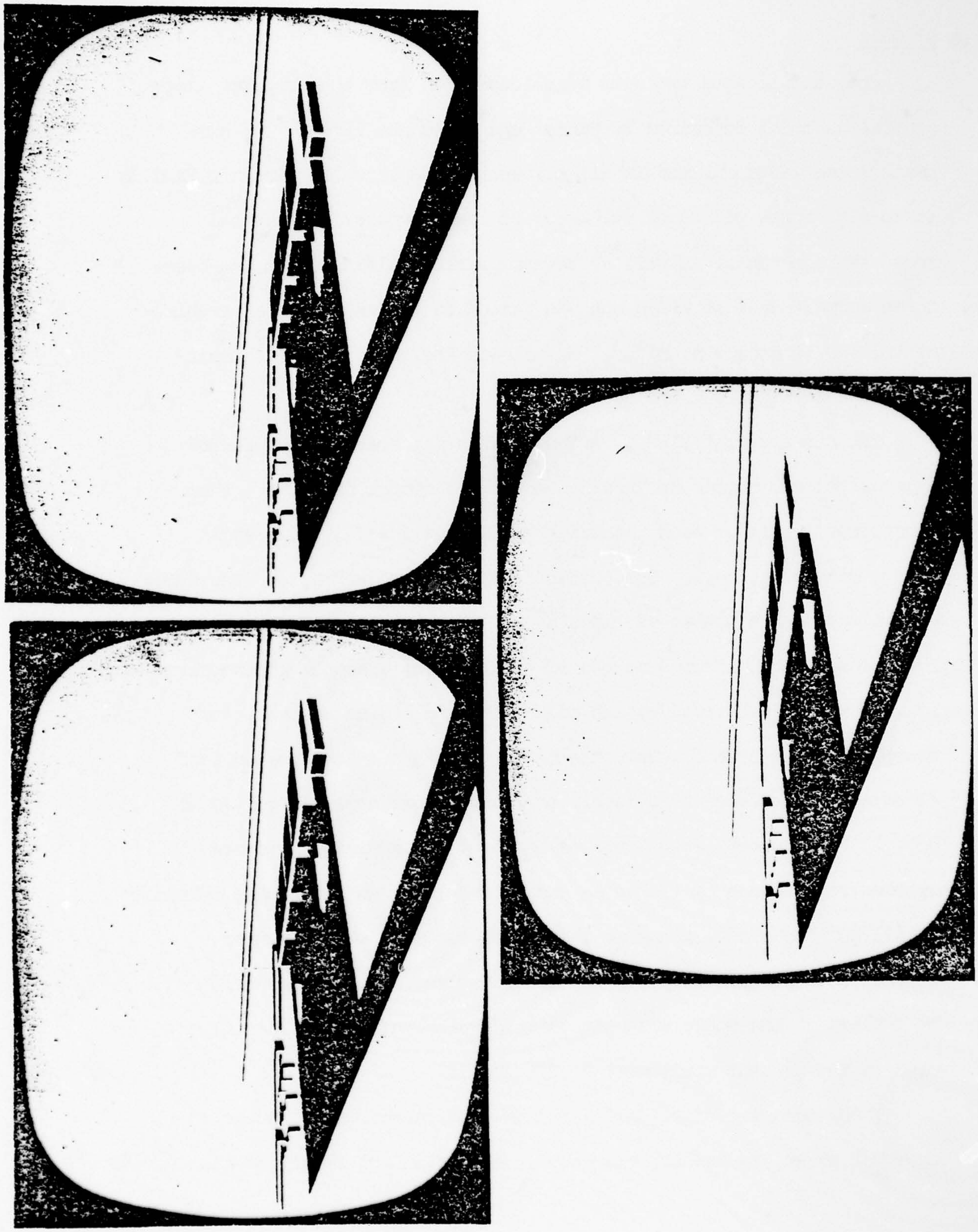


Figure 1. Stimuli used in Study I.

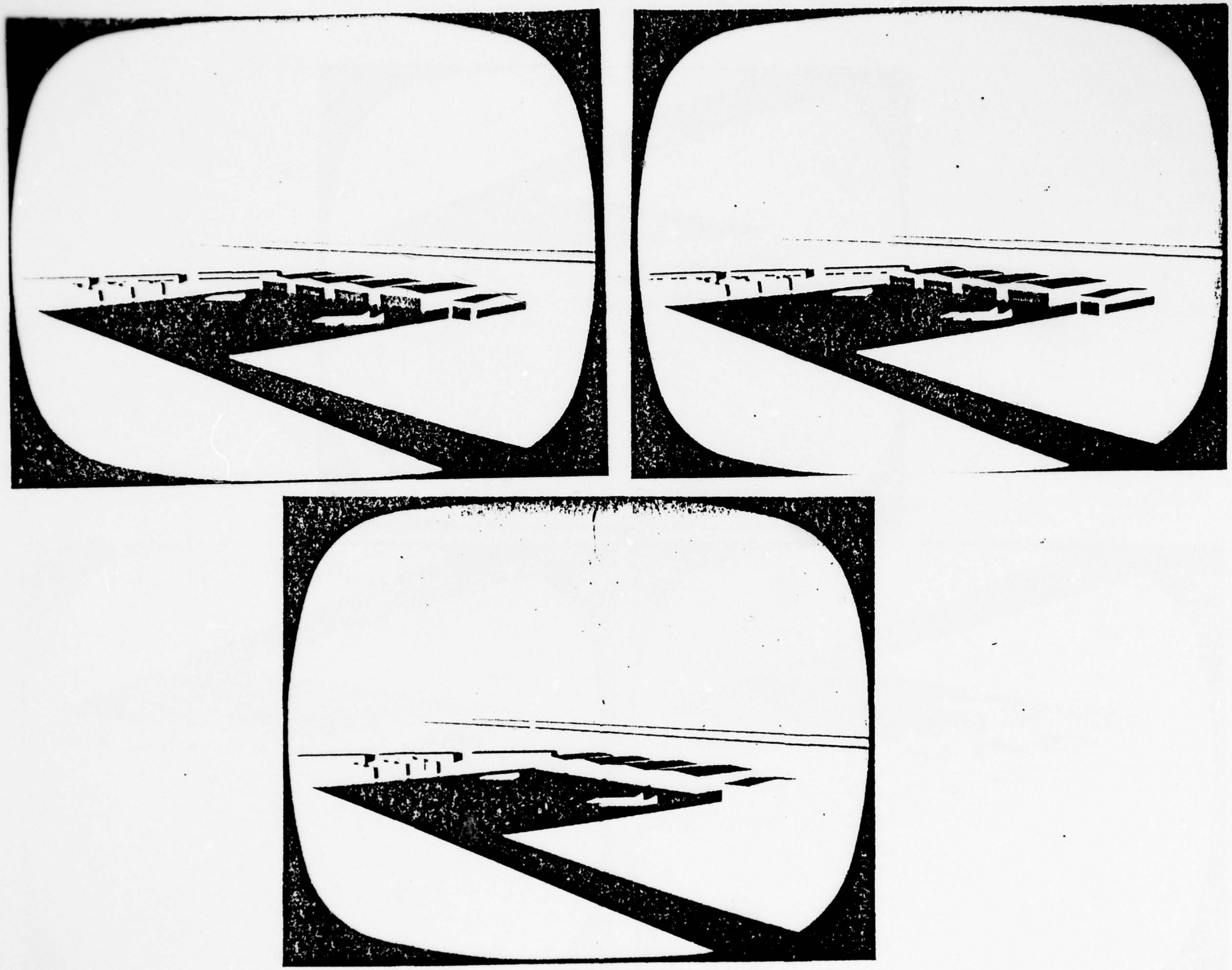


Figure 2. Stimuli used in Study I.

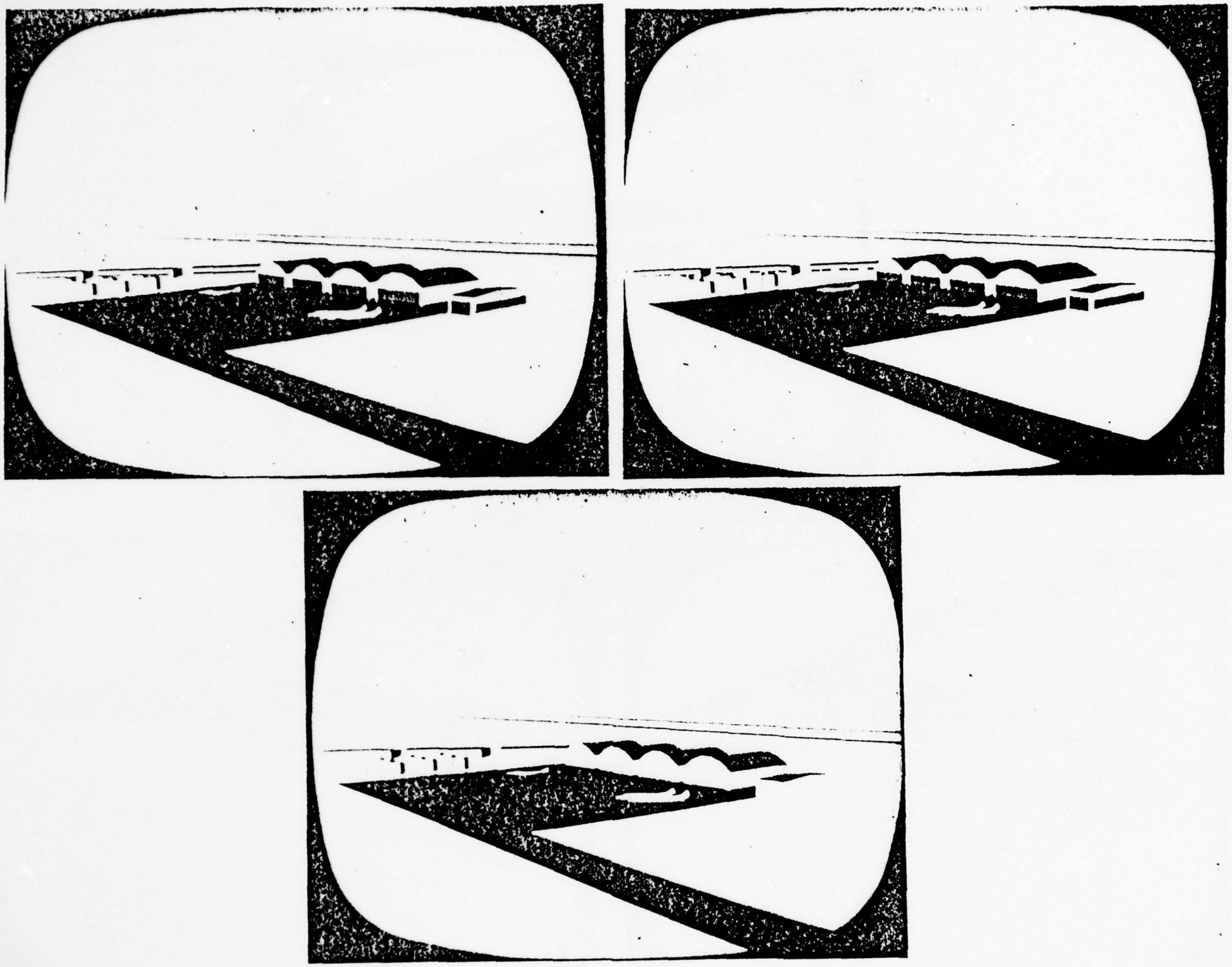


Figure 3. Stimuli used in Study I.

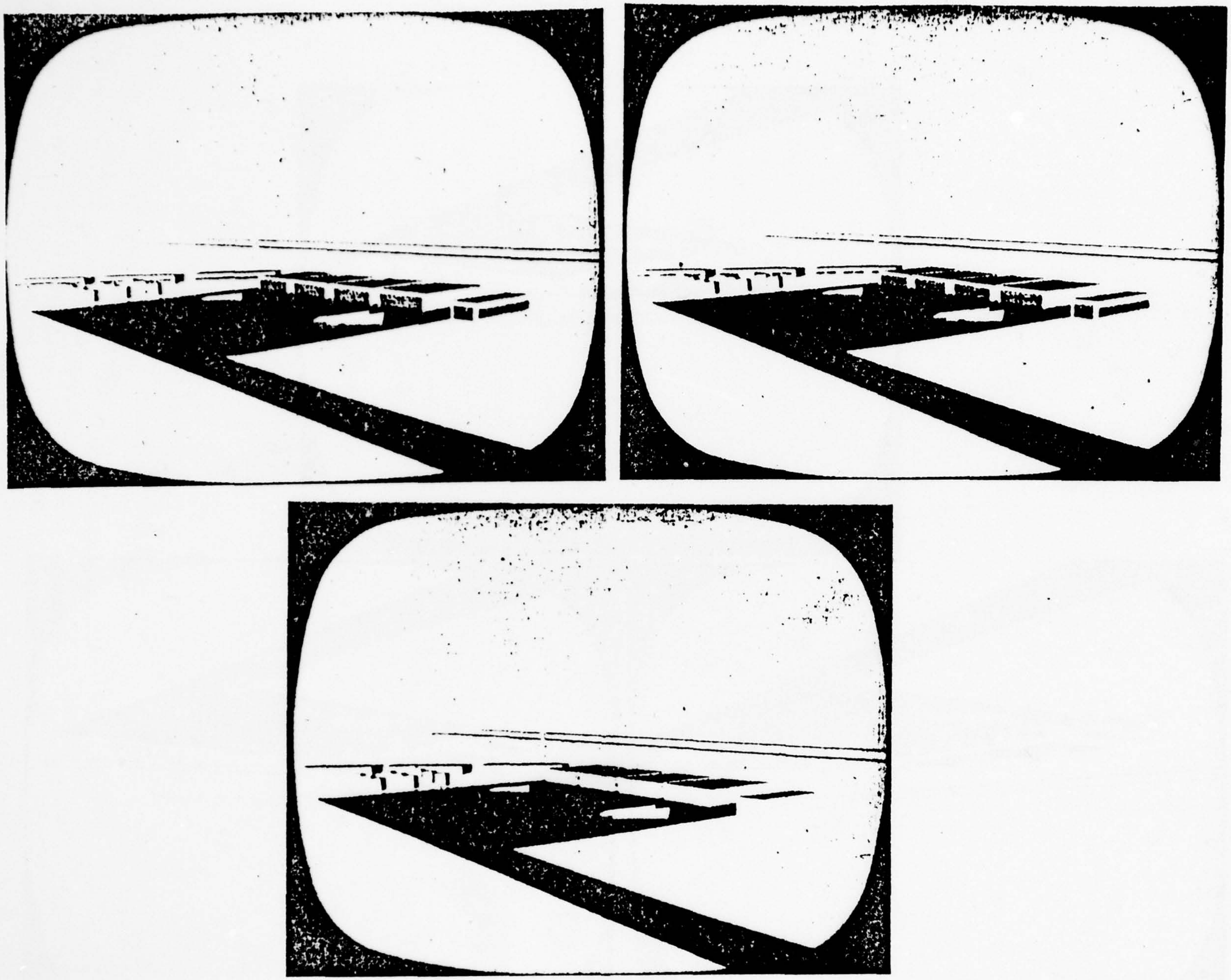


Figure 4. Stimuli used in Study I.

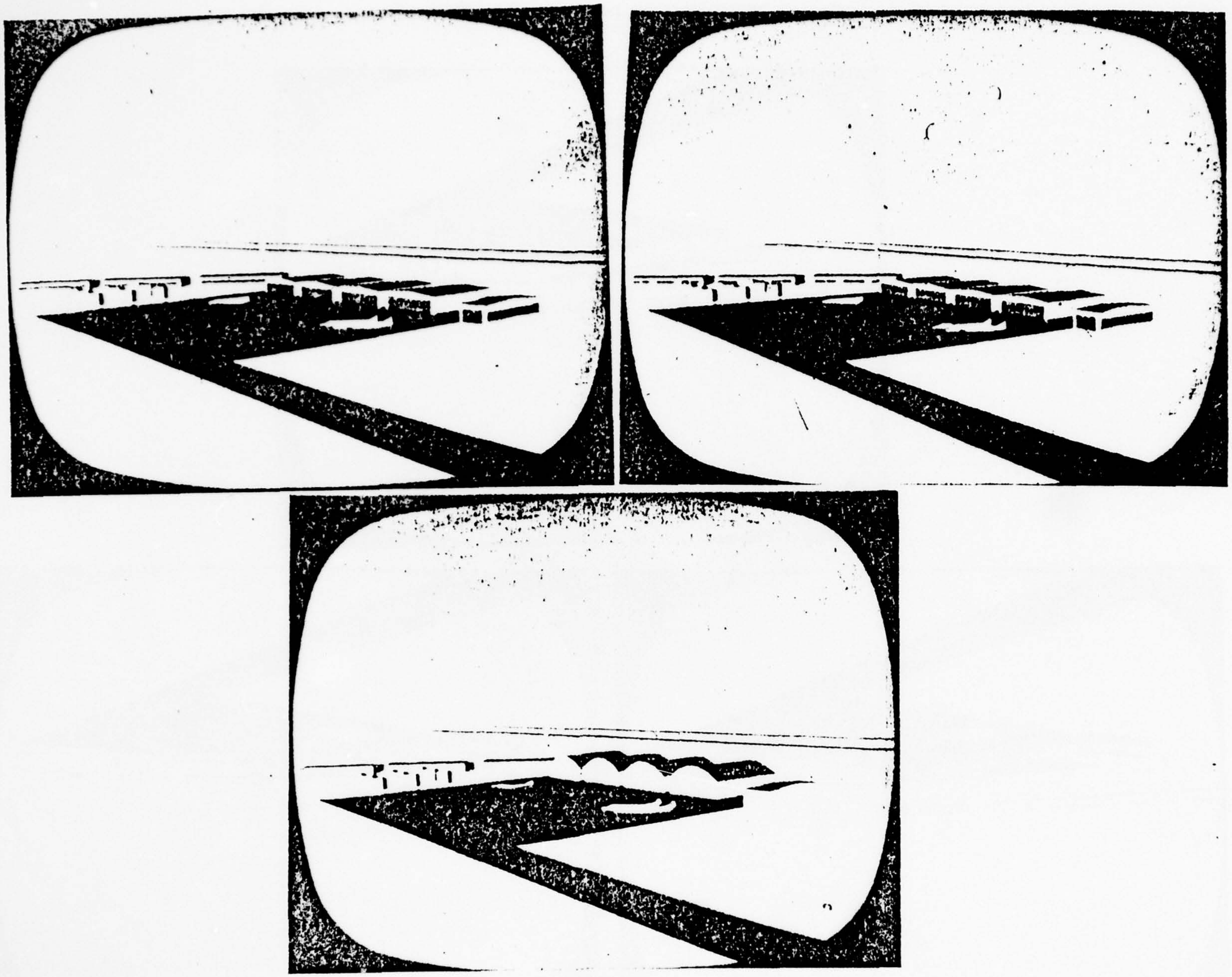


Figure 5. Stimuli used in Study I.

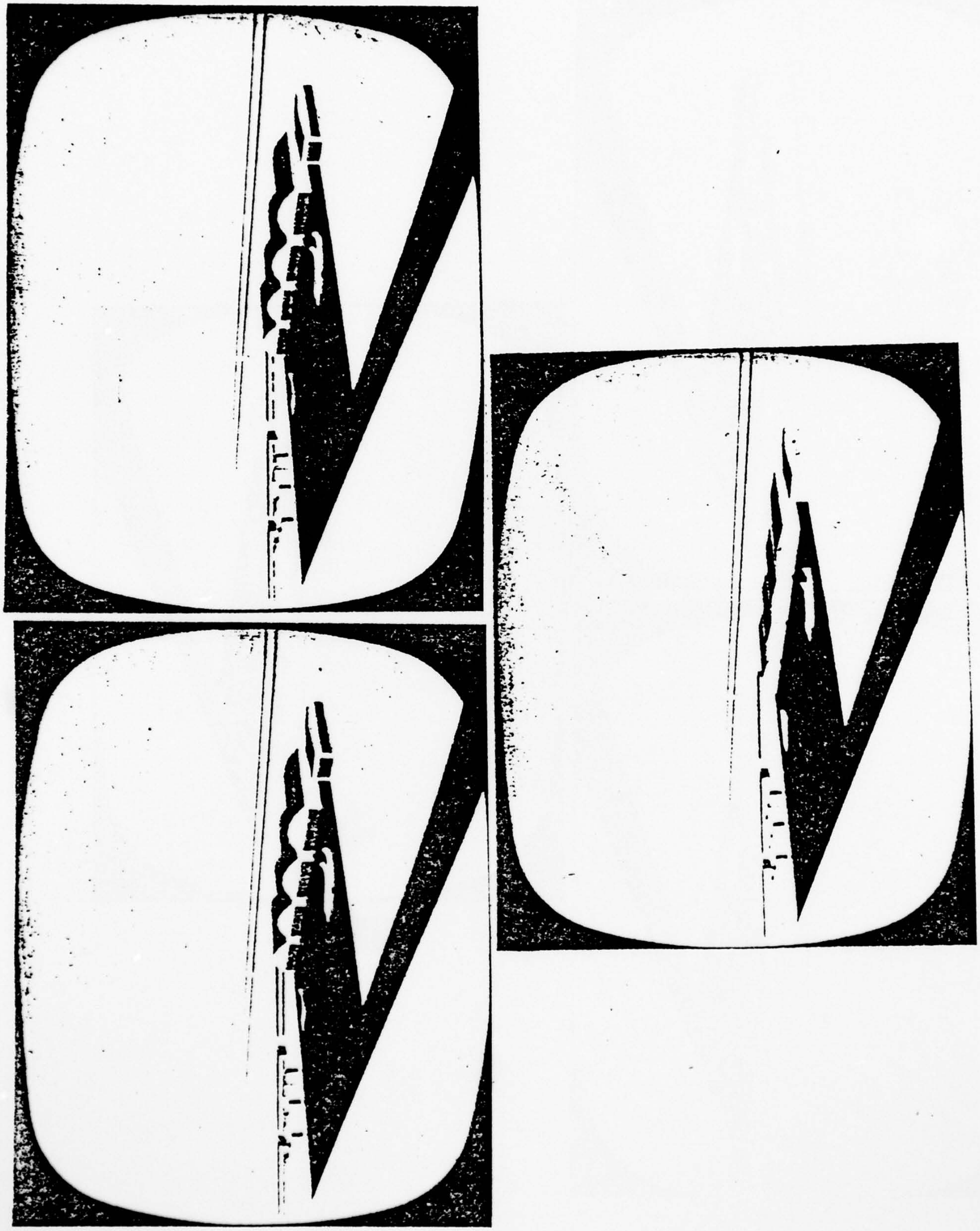


Figure 6. Stimuli used in Study I.

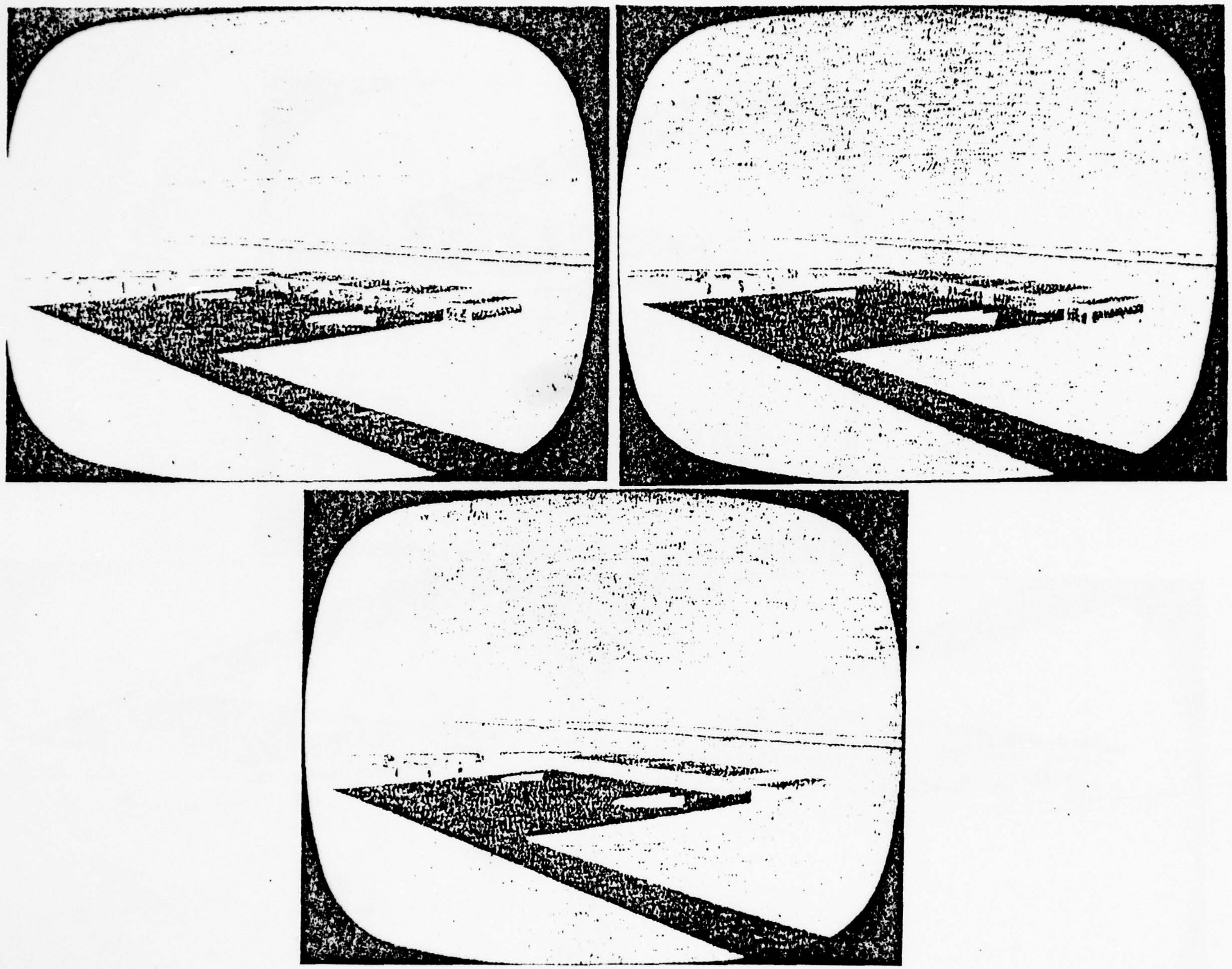


Figure 7. Stimuli used in Study I.

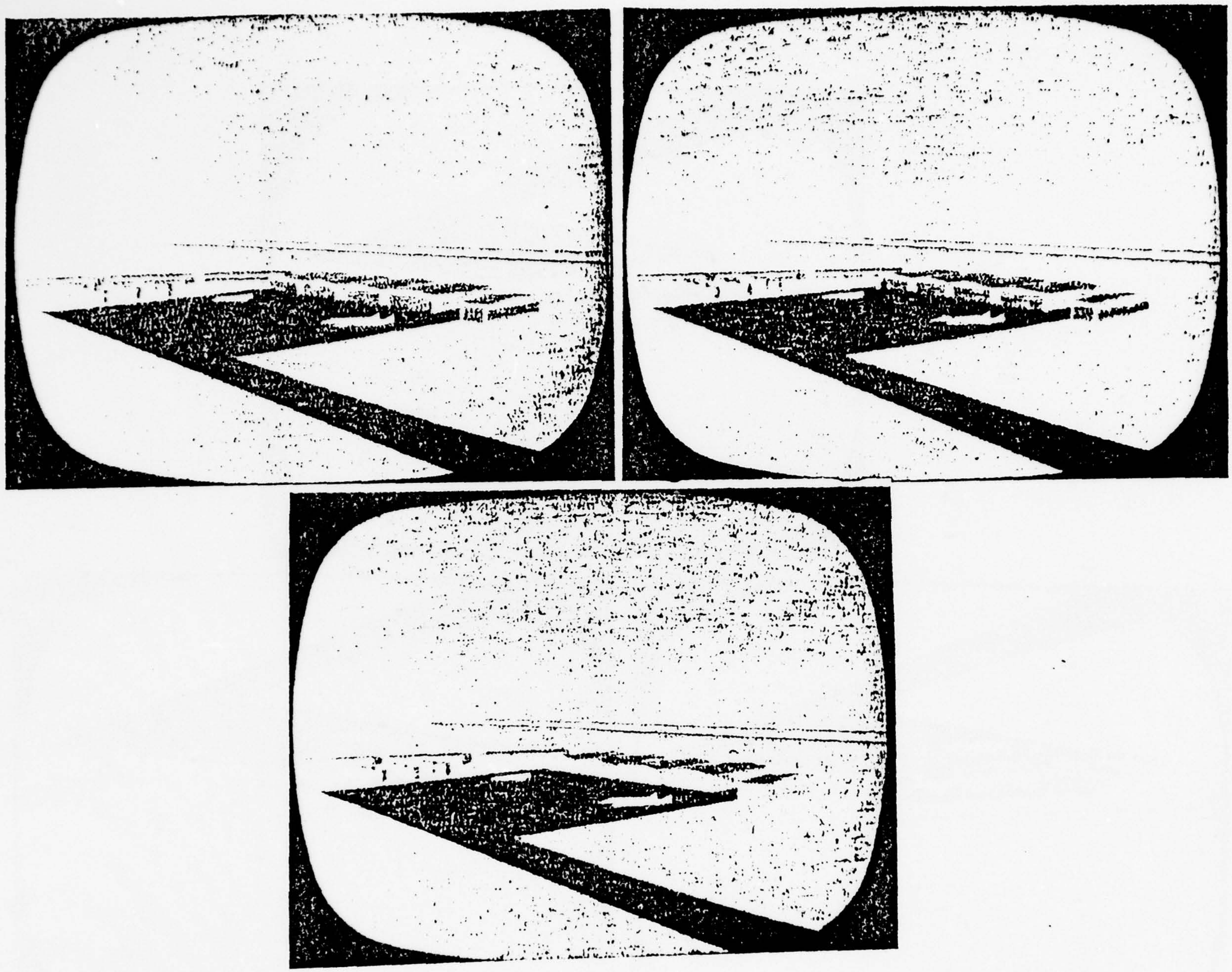


Figure 8. Stimuli used in Study I.

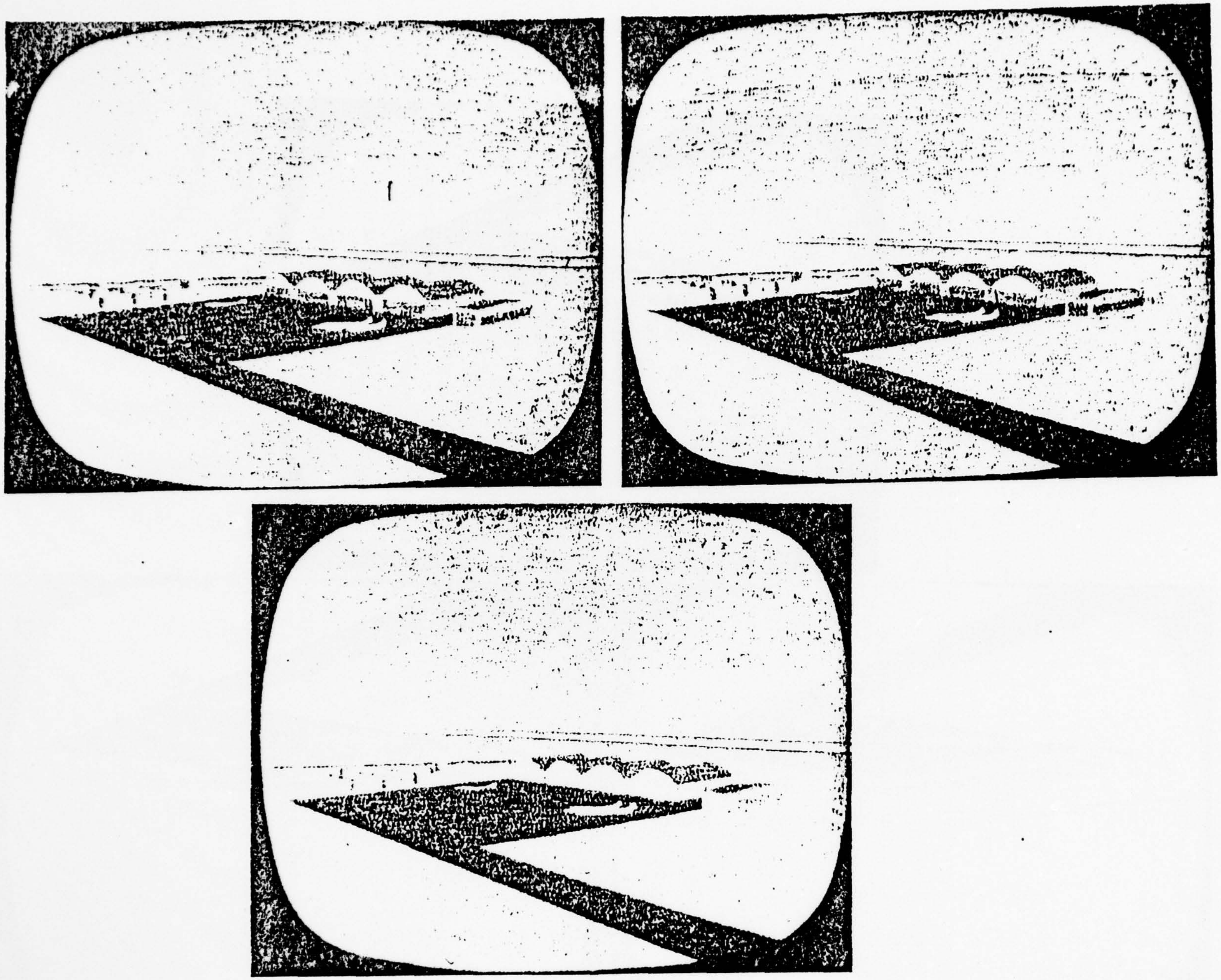


Figure 9. Stimuli used in Study I.

addition of shading to rounded figures and of a Gaussian transfer function which had a softening effect on the edges to simulate a similar effect which is ordinarily encountered in LLLTV and FLIR displays. There were two levels each of the last three parameters yielding a 3 x 2 x 2 x 2 factorial design with 24 stimuli. These stimuli are shown in Figures 10 through 17.

Procedure

The two sets of stimuli were presented consecutively to each subgroup of subjects, arranged in a different random order for each. For Group 1 and Group 2 subjects, the following instructions were read at the start of the session:

> I am going to show you a series of slides of the same scene. First, I will show you the whole series for 3 seconds each, so that you will know what they look like. Then, I will show you each slide for 15 seconds. Your job will be to rate the slides on a scale from "extremely simple" to "extremely complex" as indicated on the form provided. Do this by marking the space on the form opposite the number of the slide being shown. I will call out the slide numbers as we go. The slides are very similar to one another, but do the best you can to give your own impression of their complexity. Make your own judgements and define complexity in your own way. Are there any questions?

The slides were then shown as described and subjects rated them on a prepared form. When the ratings on the second set of slides were complete, subjects were asked to answer three questions on the back of the data sheet as follows:

1. Describe an extremely complex slide.
2. Describe an extremely simple slide.
3. What is the difference between a simple and a complex slide?

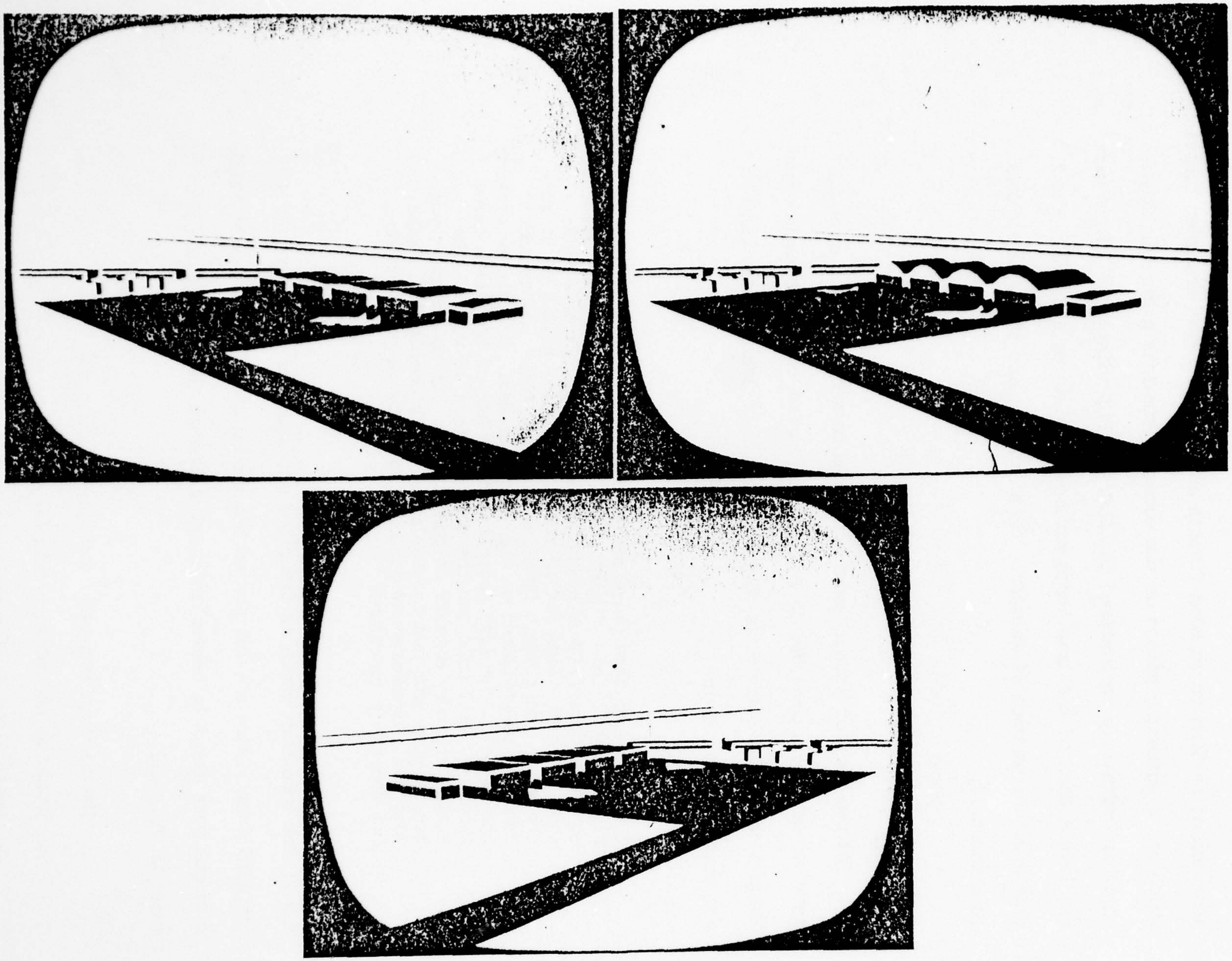


Figure 10. Stimuli used in Study II.

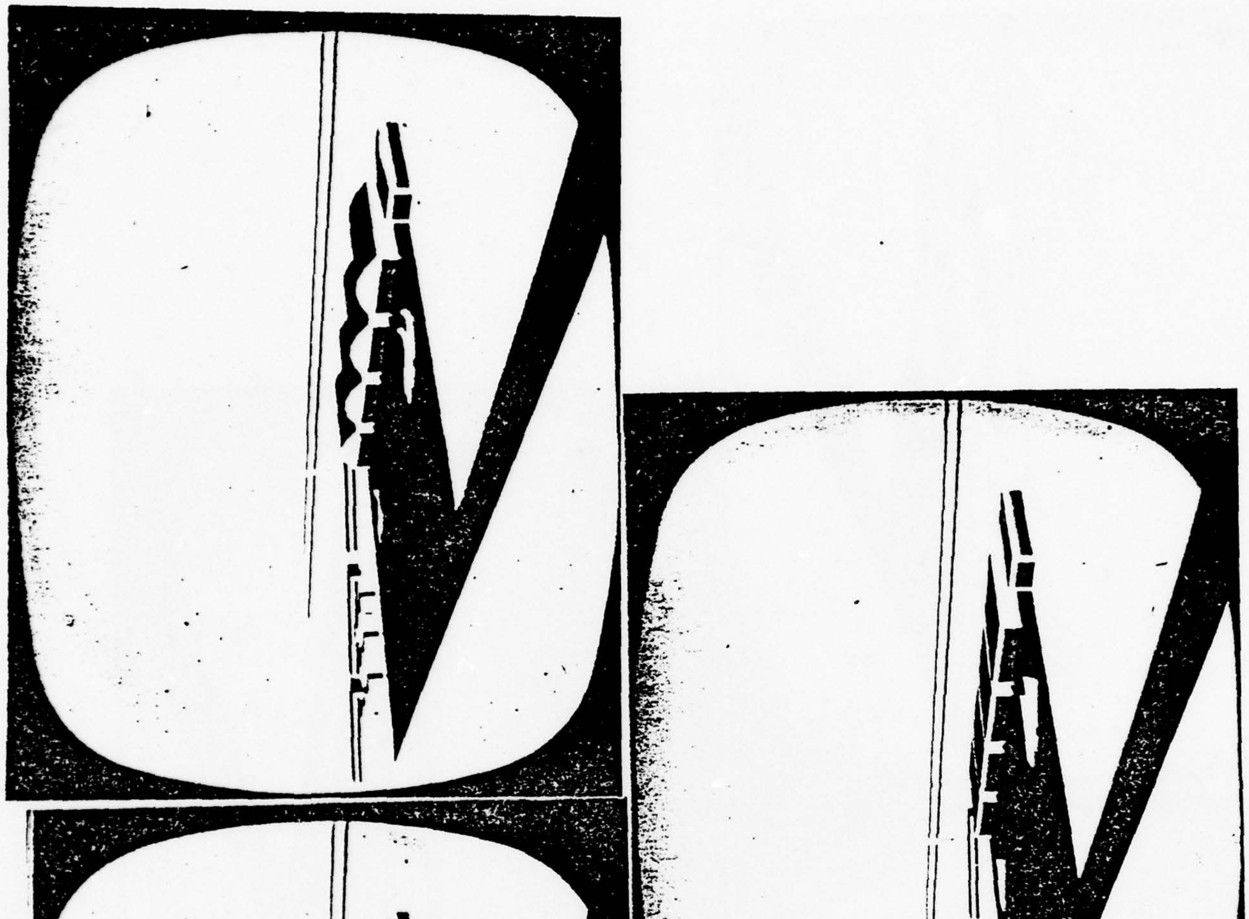


Figure 11. Stimuli used in Study II.

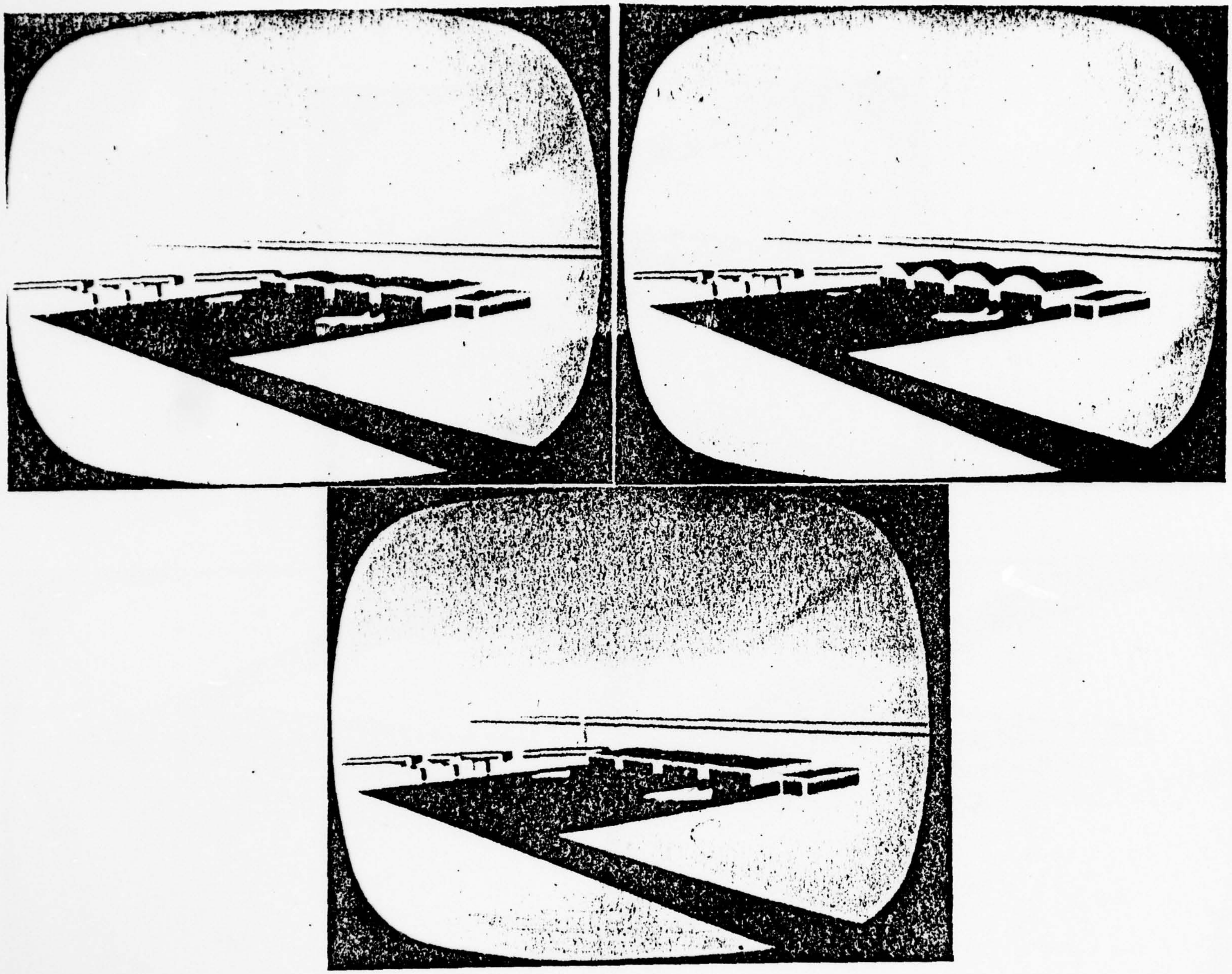


Figure 12. Stimuli used in Study II.

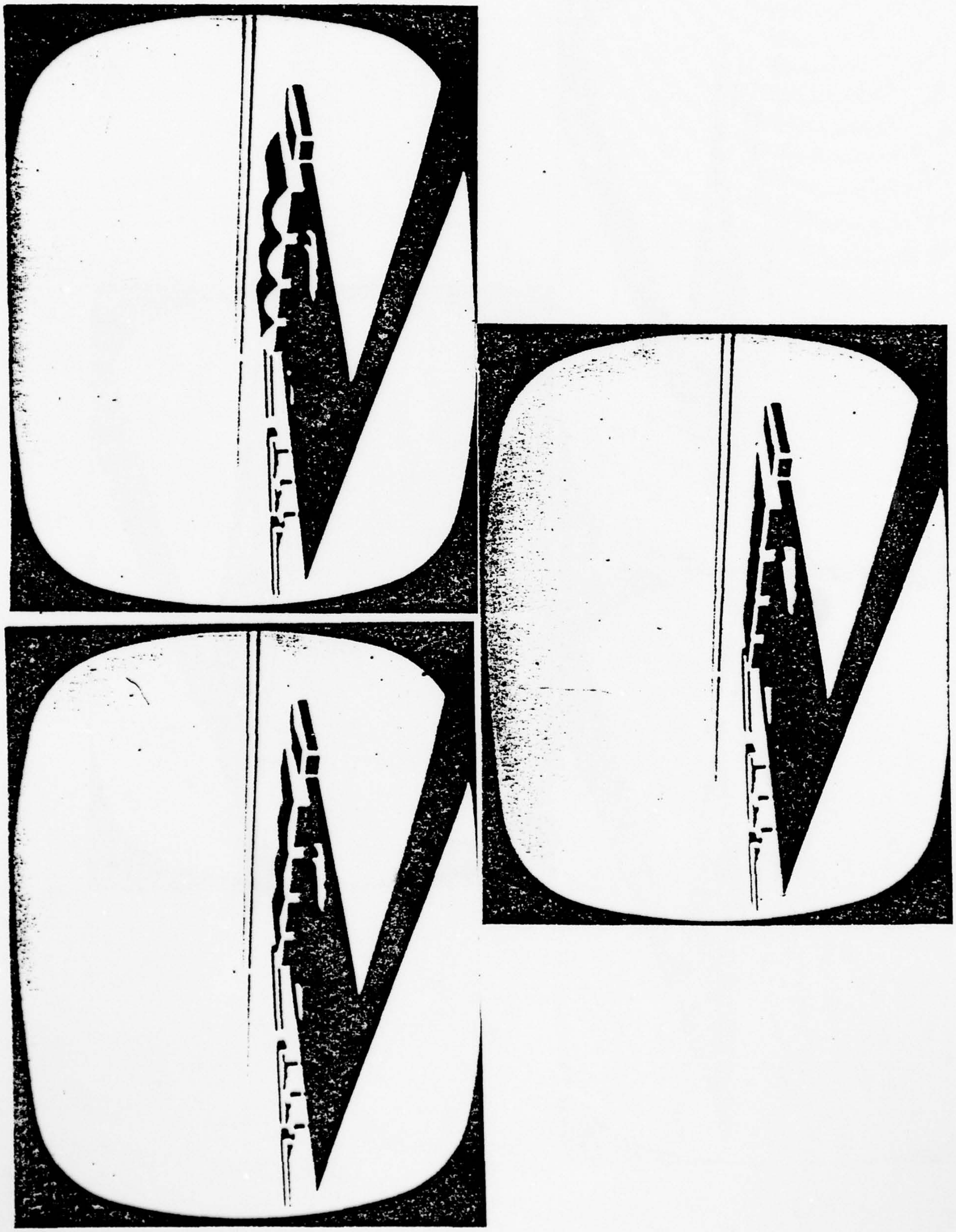


Figure 13. Stimuli used in Study II.

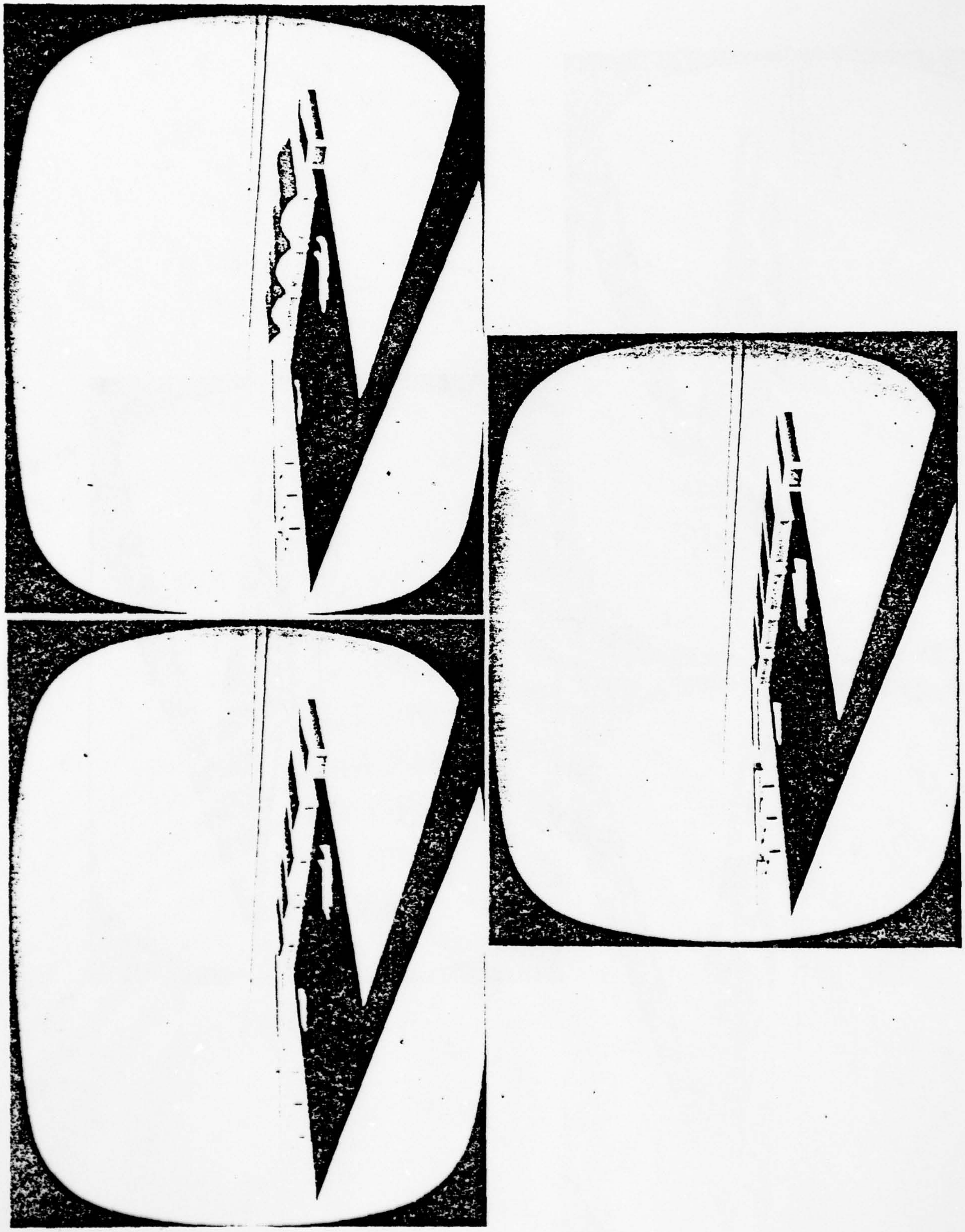


Figure 14. Stimuli used in Study II.

AD-A065 651 OHIO STATE UNIV RESEARCH FOUNDATION COLUMBUS F/G 1/3
USAF-ASEE (1978) SUMMER FACULTY RESEARCH PROGRAM (WPAFB). VOLUM--ETC(U)
NOV 78 C D BAILEY F44620-76-C-0052
UNCLASSIFIED AFOSR-TR-79-0232 NL

3 OF 5
AD A065651

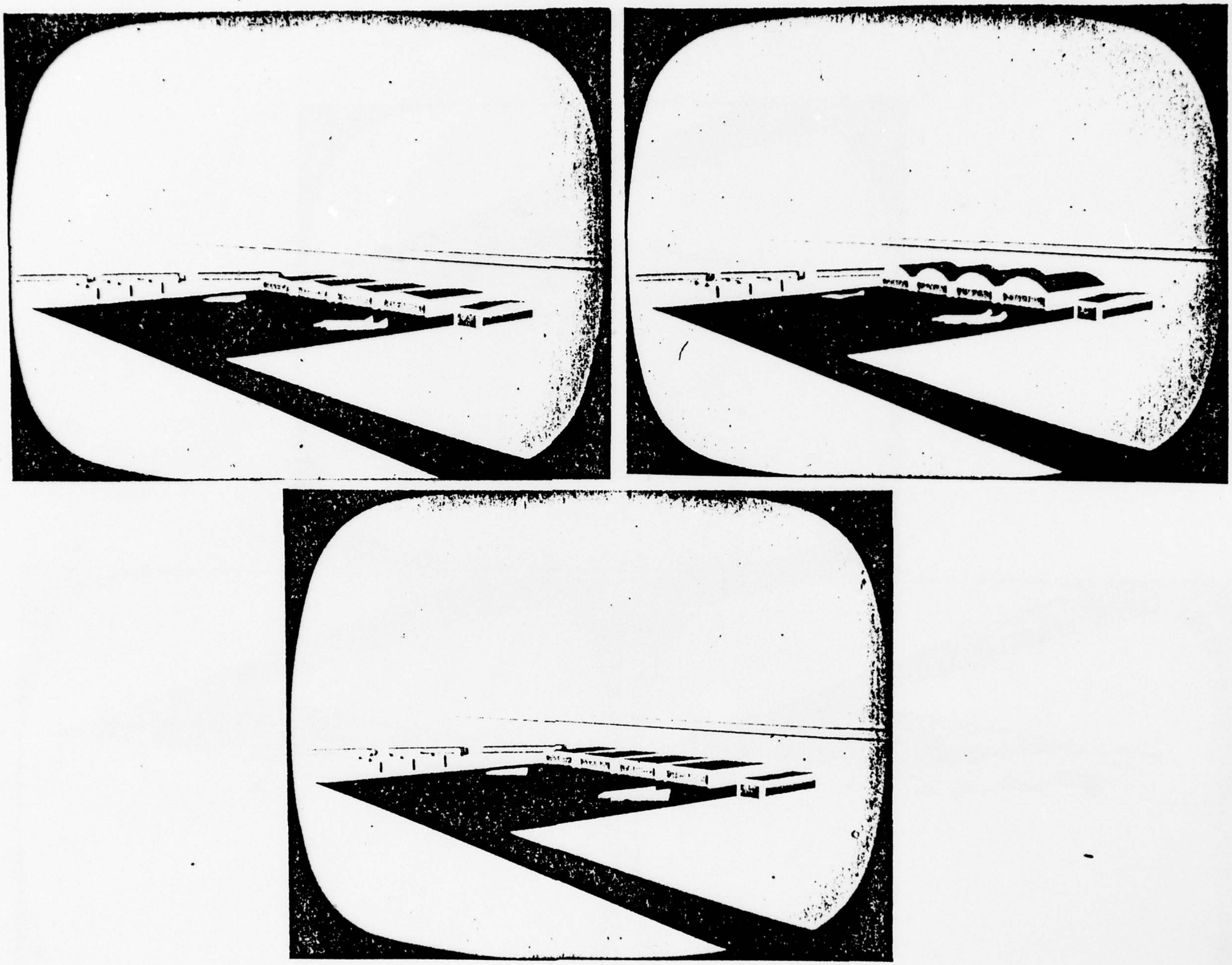


Figure 15. Stimuli used in Study II.

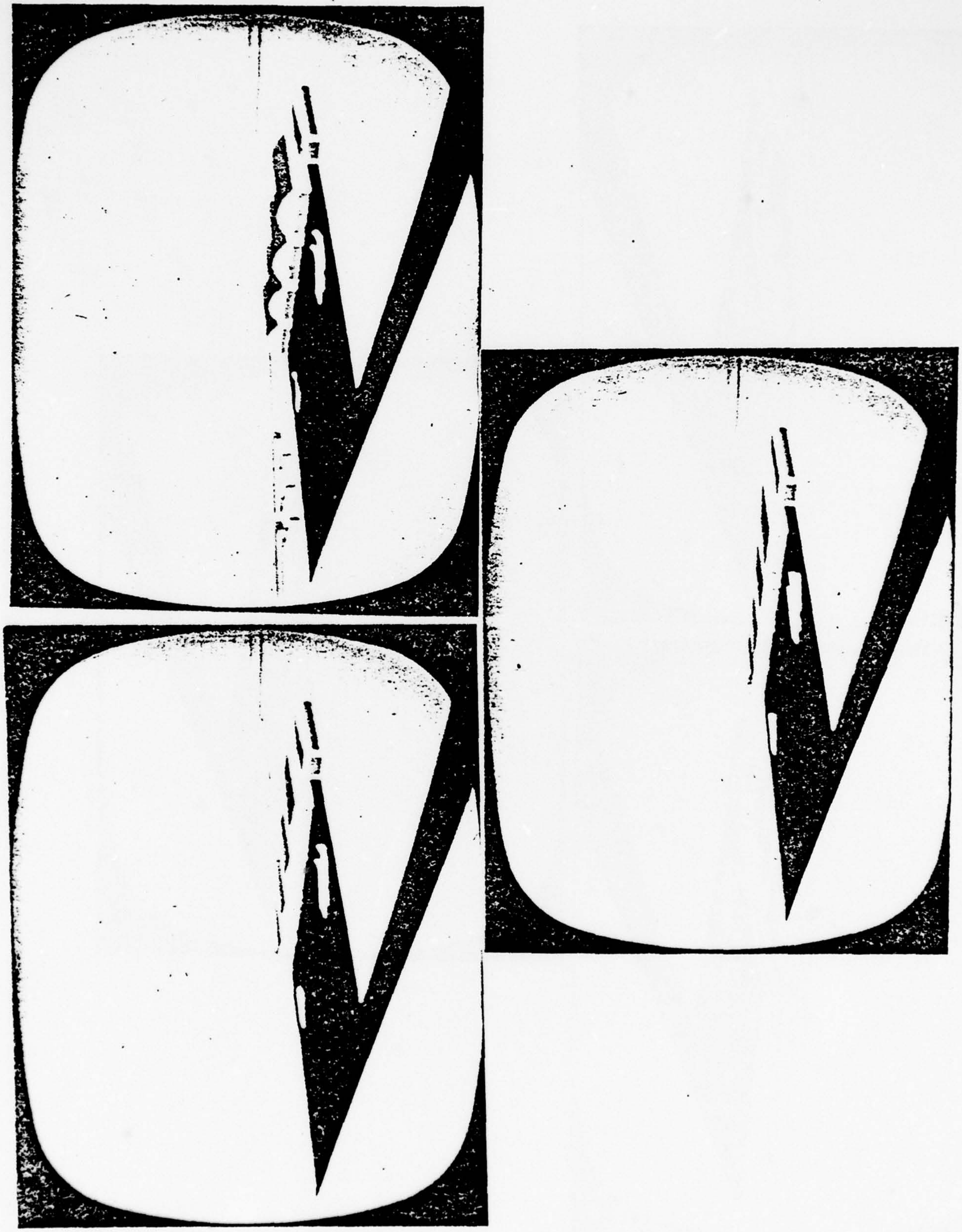


Figure 16. Stimuli used in Study II.

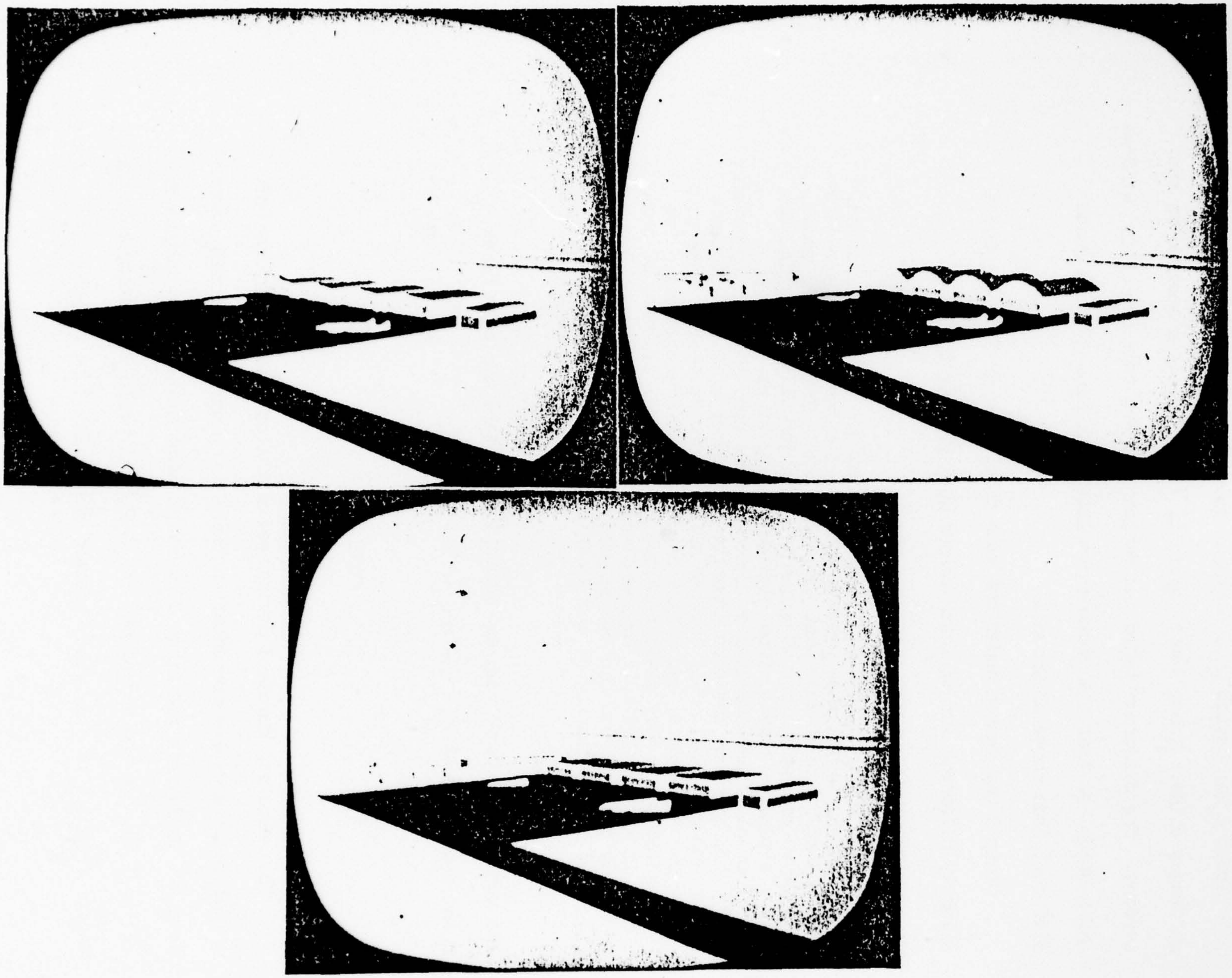


Figure 17. Stimuli used in Study II.

Group 1 was tested in a darkened room with one window covered by an opaque shade. Group 2 was tested in a regular class meeting in the evening. The classroom faced the setting sun and only one of two windows had a shade, so that the viewing conditions for Group 2 were quite different from those of Group 1.

Group 3 was tested under the same conditions as Group 1, but a slightly different set of instructions was read, as follows:

> I am going to show you a series of slides of the same scene. First, I will show you the whole series for 2 seconds each so that you will know what they look like. Then, I will show you each slide for 15 seconds. Your job will be to rate the slides on a scale from "extremely simple" to "extremely complex" as indicated on the form provided. Do this by marking the space on the form opposite the number of the slide being shown. I will call out the slide number as we go. Make your own judgements, and define complexity in your own way. The slides are very similar to one another and, for some of them, it may be difficult to see clearly defined outlines. However, do the best you can basing your judgements on your own estimate of the complexity of the scene itself. Are there any questions?

The session then proceeded as Groups 1 and 2, and the same three questions were asked at the end of the session.

RESULTS

The means and standard deviations of the complexity ratings for all of the scenes are presented in Table 2. Complexity ratings ranged from 2.3 to 5.2 for the first 27 stimuli, with standard deviations from .67 to 2.25; for the second set of 24 stimuli, mean complexity ratings ranged from 3.4 to 4.8, with standard deviations from .70 to 2.11.

TABLE 2

MEANS AND STANDARD DEVIATIONS OF COMPLEXITY RATINGS FOR 51 CIG SIMULATED SCENES

Noise Level	Number of EXT Edges	Number of INT Edges	Complexity Rating $\overline{X}$	Complexity Rating SD
0%	222	0	2.3	1.25
		59	2.7	1.15
		151	3.4	1.43
	266	0	2.8	1.23
		59	3.2	1.03
		151	3.8	1.87
	414	0	2.6	1.43
		59	3.2	1.48
		151	3.8	1.99
25%	222	0	3.1	1.20
		59	3.6	1.17
		151	4.7	1.25
	266	0	3.7	.67
		59	3.9	.88
		151	4.1	.99
	414	0	3.1	.99
		59	4.1	1.20
		151	4.6	1.17
50%	222	0	3.8	2.25
		59	4.5	1.27
		151	4.6	1.43
	266	0	3.7	1.83
		59	4.6	1.58
		151	5.2	1.56
	414	0	4.0	1.88
		59	4.6	1.43
		151	5.2	1.03

Contrast	Transfer Function	Shading	Number of Edges	Complexity Rating $\overline{X}$	Complexity Rating SD
Normal	Absent	Absent	222	3.4	1.65
			266	3.7	1.48
			414	3.4	2.11
		Present	222	3.4	1.35
			266	3.7	.82
			414	3.8	1.54
	Present	Absent	222	3.5	.85
			266	4.0	.74
			414	3.8	1.25
		Present	222	3.7	1.42
			266	4.0	.94
			414	4.3	.82
Low	Absent	Absent	222	3.9	1.45
			266	3.9	.74
			414	4.0	1.33
		Present	222	3.4	.70
			266	3.9	.82
			414	4.2	1.45
	Present	Absent	222	4.3	1.34
			266	4.0	1.05
			414	4.8	1.03
		Present	222	3.8	1.32
			266	4.3	1.49
			414	4.0	1.49

Study I

The first set of 27 scenes were analyzed separately, as Study I, and each group was also treated separately. A 3 (number of external edges) x 3 (number of internal edges) x 3 (amount of noise) factorial design with repeated measures on the same subjects (Kirk, 1968) was used to examine the data separately for each group of students. The complexity rating per subject per cell was used as the sole dependent variable.

The results of an analysis of variance for Group 1 are presented in Table 3 where it can be seen that there are no significant effects. However, a graph of the effect of internal edges on complexity ratings for the three noise levels shown in Figure 18 seemed to indicate a possible interaction between these two variables while an examination of the raw data led to the observations that low variability due to internal edges might have obscured this interaction, and that the effect of internal edges was overshadowed by that of noise. Therefore, two further analyses were undertaken. The first was a Friedman two-way analysis of variance by ranks (Siegel, 1956) which showed a significant interaction between noise level and internal edges ($X^2 = .006$, $p < .001$). An examination of Figure 18 shows that, for the no-noise condition, complexity ratings rise with increasing number of internal edges but this effect drops out when noise is added to the display. Because the effect of internal edges seemed to operate only in the no-noise condition, a separate analysis of variance was performed on the complexity ratings for only those nine scenes in which there was no noise. As expected, the effect of internal edges was significant in the absence of noise ($F = 4.46$, $p < .05$).

TABLE 3

ANALYSIS OF VARIANCE FOR STUDY I, GROUP 1

SOURCE	SUM OF SQUARES	DF	MEAN SQUARE	F-RATIO
Betw Err	184.65	20	9.23	
Noise (A)	4.08	2	2.04	
W/N Err 1	832.05	40	20.80	
Ext Edge (B)	1.27	2	.64	
W/N Err 2	74.87	40	1.87	
Int Edge (C)	17.06	2	8.53	2.51 n.s.
W/N Err (3)	136.20	40	3.40	
AB	1.16	4	.29	
W/N Err 4	64.47	80	.80	
AC	4.90	4	1.22	1.83 n.s.
W/N Err 5	53.62	80	.67	
BC	.73	4	.18	
W/N Err 6	60.46	80	.75	
ABC	4.43	8	.55	
W/N Err 7	94.61	160	.59	
TOTAL	1534.58	566	2.71	

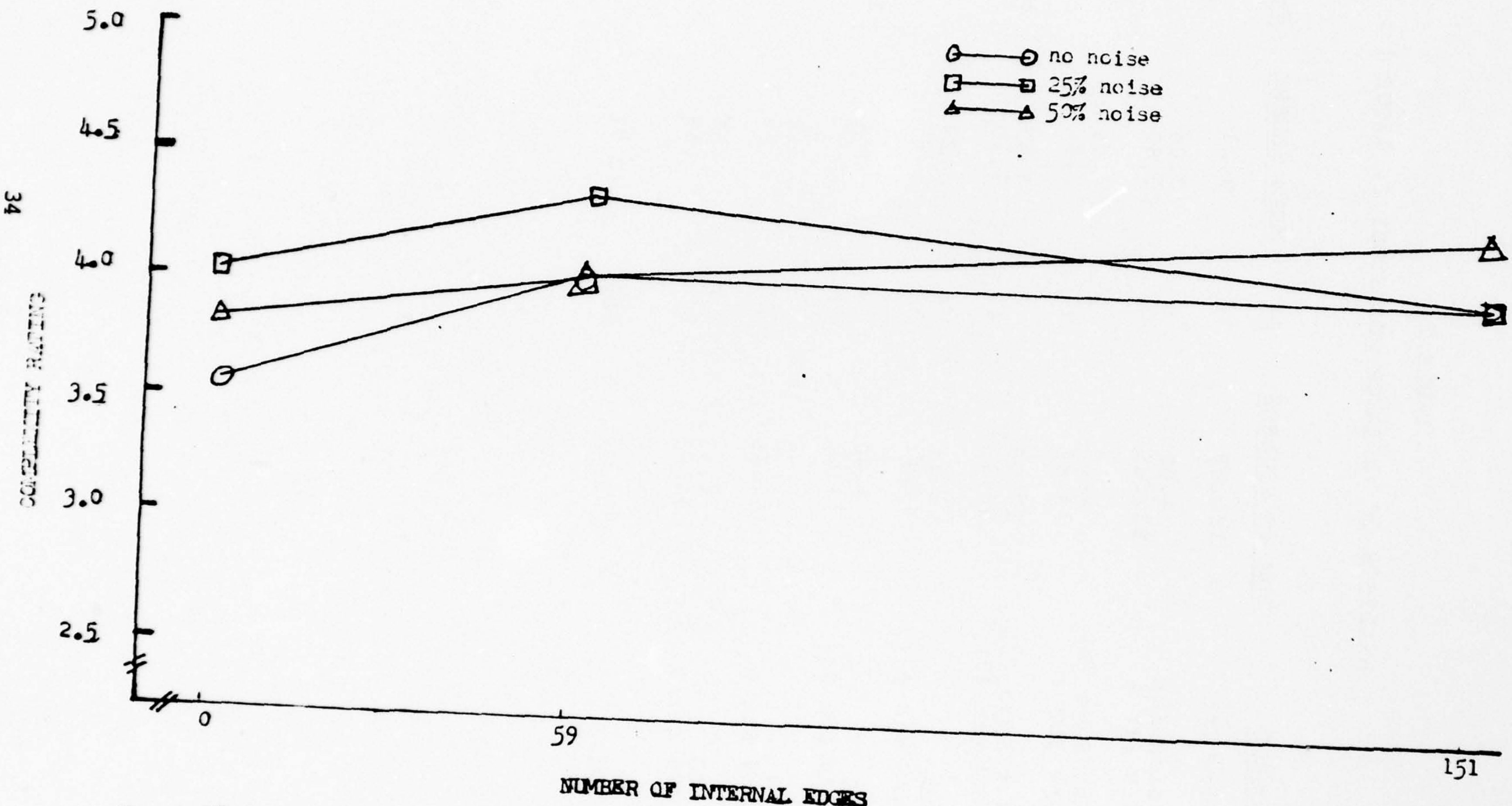


Figure 18. The effect of number of internal edges on complexity ratings for three noise levels for Group 1.

The results of a similar analysis for Group 2 are presented in Table 4. The only significant effect here is that of noise level. The effect of noise on complexity ratings is seen clearly in Figure 19 where the interaction with internal edges is no longer evident. Instead, noise is the sole determiner of complexity rating with increased noise leading to an increase in complexity rating. This is the group which was tested under a condition of relatively high ambient illumination. It would seem as if the lowered contrast occasioned by this condition accentuated the effect of noise to the point where it overshadowed all other effects. Internal evidence from Study II which supports this suggestion is presented later.

It was clear from the above analysis that noise itself was considered by subjects to be a primary factor in the judgement of complexity. Moreover, there was evidence in the raw data and in the answers to the questions asked at the end of the sessions that while noise increased complexity over all subjects, there was a substantial minority for whom its masking effect decreased complexity. Since noise had originally been introduced in order to determine whether or not it interfered with the effects of edges and it was clear that many subjects were interpreting it as an integral part of the scene, it was decided to run a third group of subjects with the slightly altered instructions reported above in order to insure more uniform interpretation of noise. An analysis of variance was performed as above on the resulting data, and is presented in Table 5. Noise is still a significant variable but the main effect of number of internal edges is also significant. Number of external edges had no significant effect. The effect of noise and internal edges is shown in Figure 20. It can be seen that complexity ratings increase with increasing numbers of edges and also with increasing amounts of

TABLE 4

ANALYSIS OF VARIANCE FOR STUDY I, GROUP 2

SOURCE	SUM OF SQUARES	DF	MEAN SQUARE	F-RATIO	PROB
Betw Err	140.87	15	9.39		
Noise (A)	104.09	2	52.04	11.41	<0.001
W/N Err 1	136.88	30	4.56		
Ext Edge (B)	.75	2	.38		
W/N Err 2	33.32	30	1.11		
Int Edge (C)	.92	2	.46		
W/N Err 3	38.71	30	1.29		
AB	1.48	4	.37		
W/N Err 4	32.67	60	.54		
AC	1.40	4	.35		
W/N Err 5	35.19	60	.59		
BC	1.06	4	.27		
W/N Err 6	18.42	60	.31		
ABC	2.32	8	.29		
W/N Err 7	50.64	120	.42		
TOTAL	598.72	431	1.39		

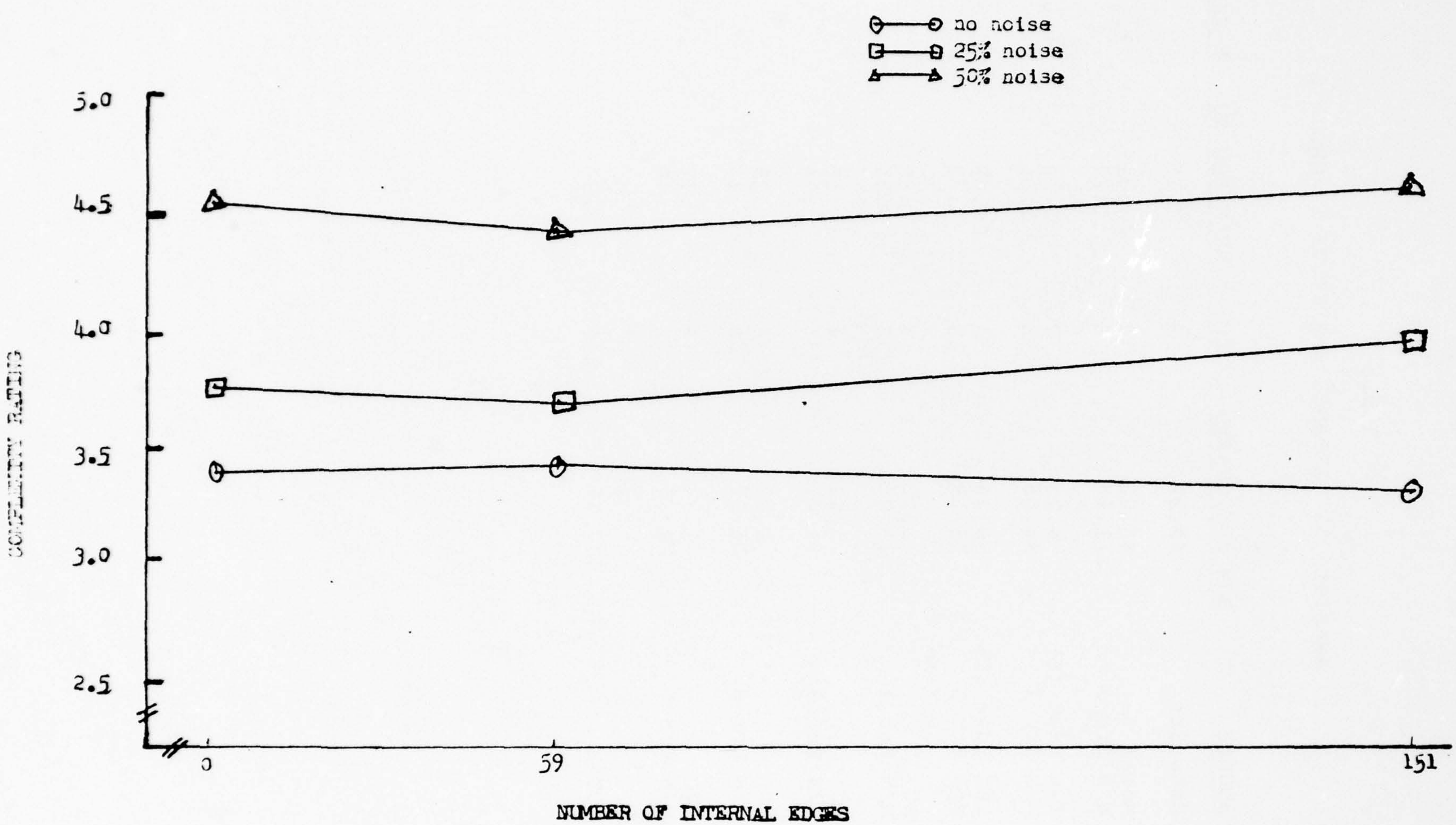


Figure 19. The effect of number of internal edges on complexity ratings for three noise levels for Group 2.

TABLE 5

ANALYSIS OF VARIANCE FOR STUDY I, GROUP 3

SOURCE	SUM OF SQUARES	DF	MEAN SQUARE	F-RATIO	PROB
Betw Err	65.22	9	7.25		
Noise (A)	86.02	2	43.01	3.66	< .04
W/N Err 1	211.24	18	11.74		
Ext Edge (B)	4.29	2	2.14	1.13	n.s.
W/N Err 2	34.08	18	1.89		
Int Edge (C)	58.96	2	29.48	6.23	< .01
W/N Err 3	85.19	18	4.73		
AB	1.22	4	.30		
W/N Err 4	19.96	36	.55		
AC	.62	4	.15		
W/N Err 5	19.45	36	.54		
BC	1.02	4	.25		
W/N Err 6	15.27	36	.42		
ABC	6.13	8	.77	1.92	n.s.
W/N Err 7	28.68	72	.40		
TOTAL	637.37	269	2.37		

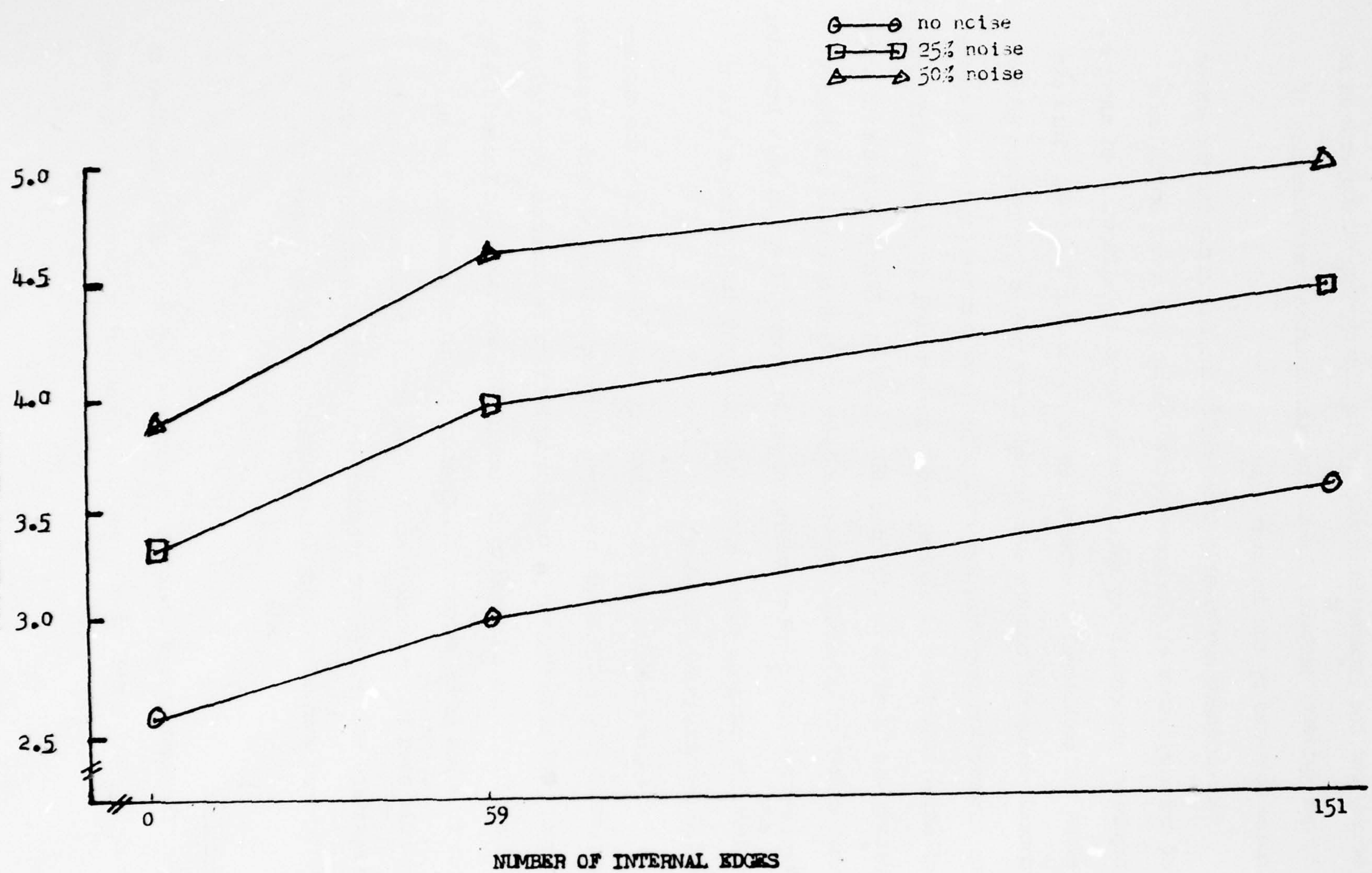


Figure 20. The effect of number of internal edges on complexity ratings for three noise levels for Group 3.

noise, and the interaction found in the first group of subjects is no longer evident, probably due to the more uniform interpretation of noise fostered by the instructions.

Since there are clearly three different functions for the effect of internal edges at the three noise levels, the noise levels were separated and complexity ratings were plotted as a function of internal edges for each level of external edges in Figures 21 to 23. This was considered useful because even though there was no significant effect or interaction involving external edges, the correlation between number of edges regardless of whether they were external or internal and complexity rating was found to be .81 ($p < .01$). In fact, Figure 21 seems to show that number of external edges does lead to slightly higher complexity ratings for the no-noise condition, while Figure 22 and 23 show irregular effects of external edges, again indicating the disruptive effects of noise on complexity judgements.

The above results are summarized in Table 6 which gives the characteristics that a CIG scene must have in order to equal a given complexity level. The distribution of complexity ratings is somewhat truncated when compared to those obtained by de Groot (1978) and LeMay and Locher (1978), but this was to be expected considering the close similarity of all the stimuli used in the present study. The effect of noise is to add to complexity but it does not interact with edges so that its effects may be ignored when considering other aspects of display complexity.

Study II

The same type of analysis was employed for Study II as was used in Study I. A 2 (level of contrast) x 2 (presence or absence of a transfer

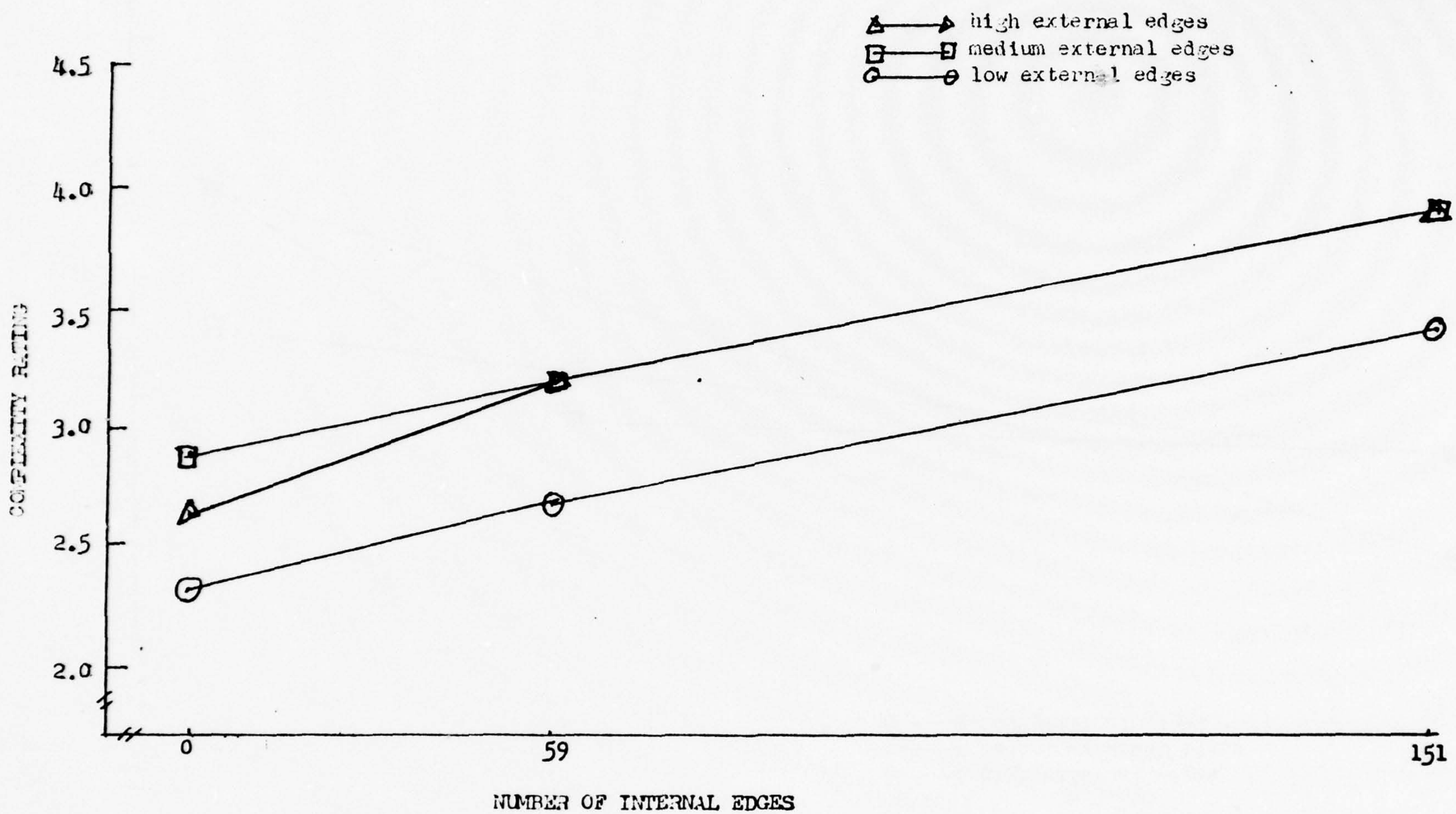


Figure 21. The effect of number of internal edges on complexity ratings at three levels of external edges under no noise condition (Group 3 data).

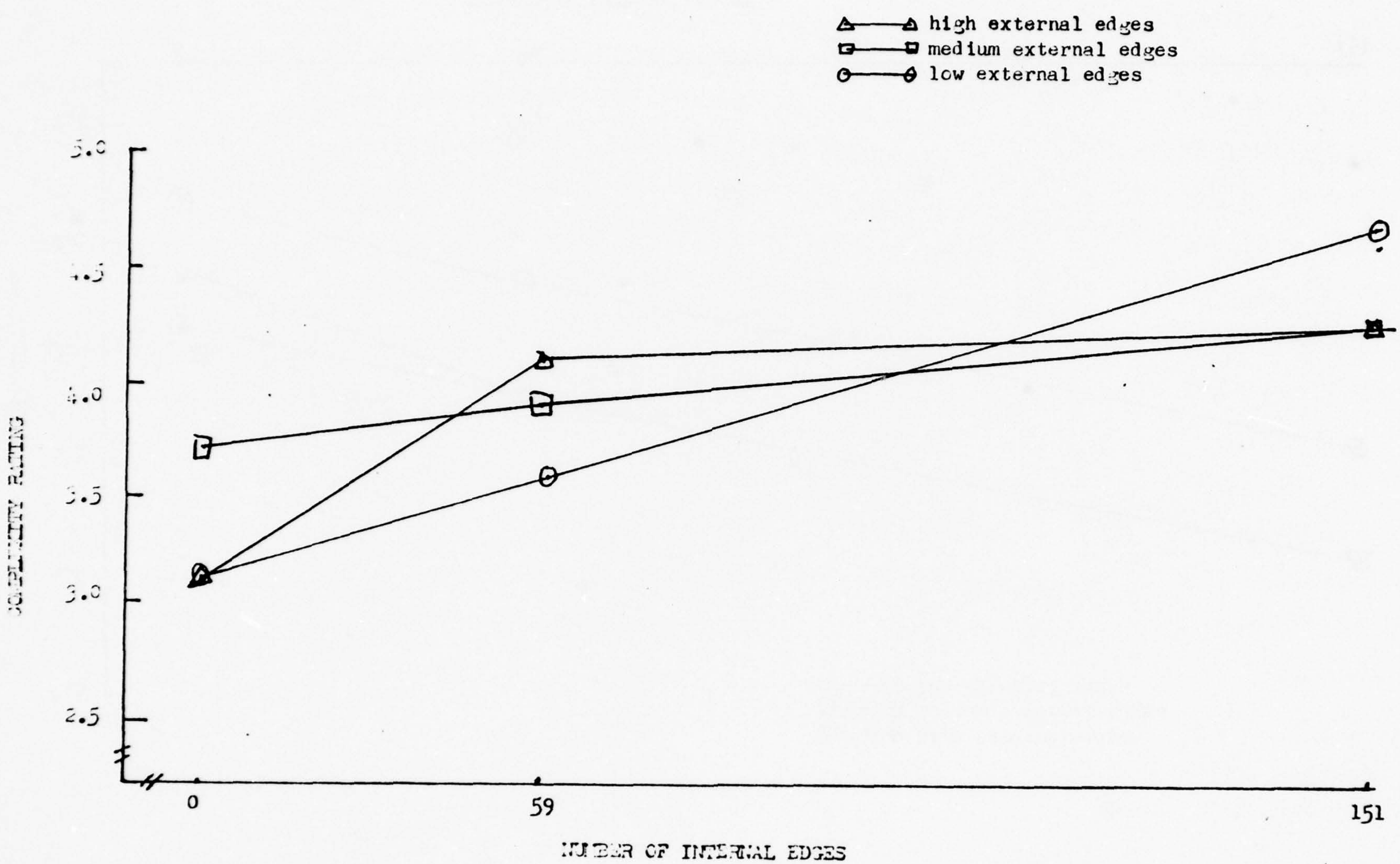


Figure 22. The effect of number of internal edges on complexity ratings at three levels of external edges under a 25% noise condition (Group 3 data)

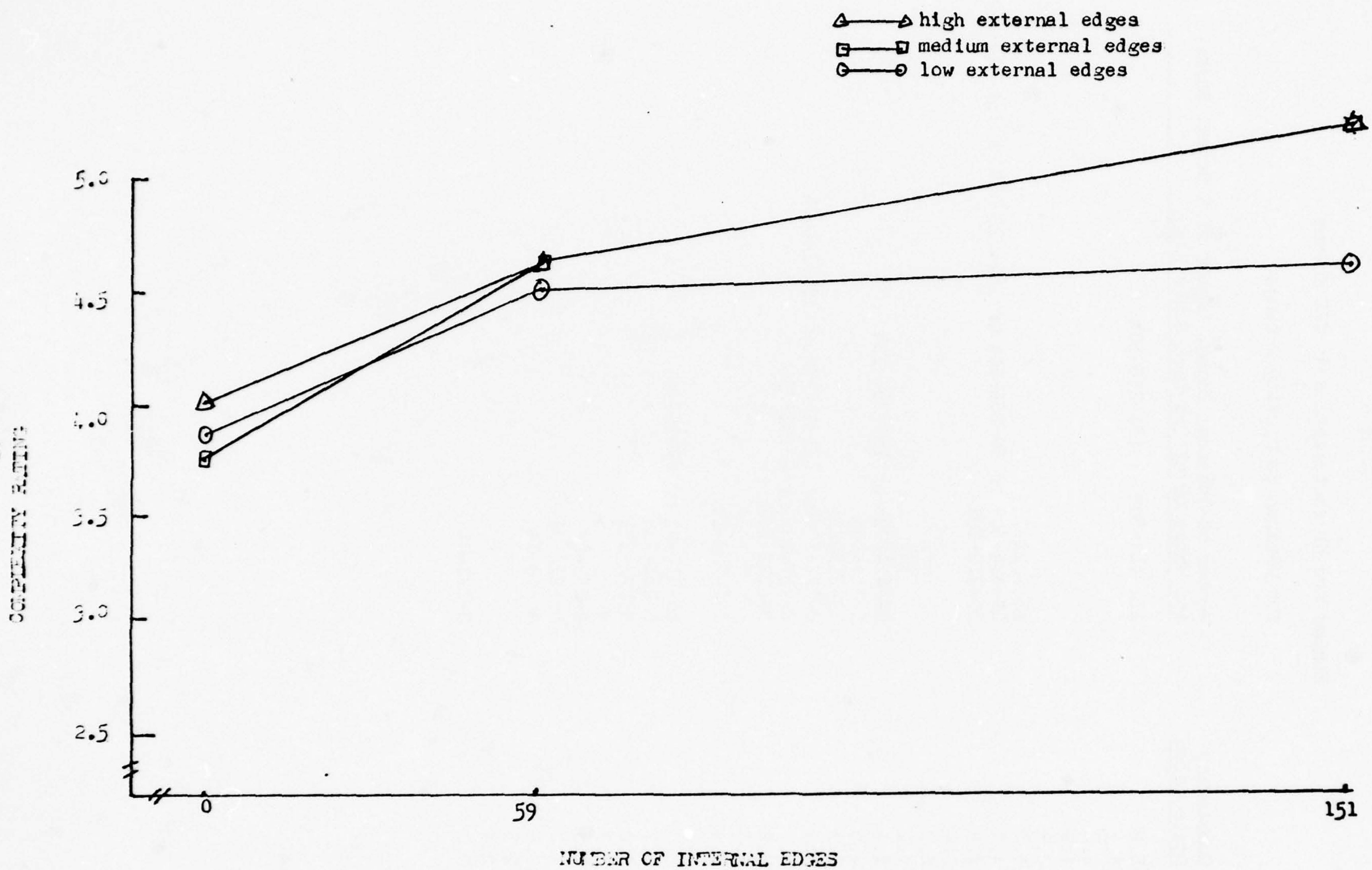


Figure 23. The effect of number of internal edges on complexity ratings at three levels of external edges under a 50% noise condition.

TABLE 6

Number and Characteristics of CIG Scenes

for Various Complexity Ratings

Complexity Scale Value	Number of Internal Edges, Number of External Edges, and Noise Level for Each Scale Value
5.2	151-414-50% 151-266-50%
5.1	
5.0	
4.9	
4.8	
4.7	151-222-25%
4.6	59-414-50% or 59-266-50% or 151-222-50% or 151-414-25%
4.5	59-222-50%
4.4	
4.3	
4.2	
4.1	59-414-25% or 151-266-25%
4.0	0-414-50%
3.9	59-266-25%
3.8	0-222-50% or 151-414-0% or 151-266-0%
3.7	0-266-50% or 0-266-25%
3.6	59-222-25%
3.5	
3.4	151-222-0%
3.3	
3.2	59-414-0% or 59-266-0%
3.1	0-414-25%
3.0	0-222-25%
2.9	
2.8	0-266-0%
2.7	59-222-0%
2.6	0-414-0%
2.5	
2.4	
2.3	0-222-0%

function) x 2 (presence or absence of shading) x 3 (number of edges) design resulted. A summary of the analysis of the data for Group 1 is shown in Table 7. The only significant factor was the third order interaction. Although such interactions are difficult to interpret, this would seem to indicate that all of the factors result in some variability but that the variability due to any one factor or combination of factors up to three is too small to produce a significant effect.

The analysis of variance for the same set of data for Group 2 is shown in Table 8. In this case, both contrast level and the presence of a transfer function show significant main effects. The fading of contrast in this case led to higher complexity judgements (it resulted in lower complexity judgements for Group 1 but this effect was not significant) as did the addition of noise in Study I. This was the group for which the level of ambient illumination was quite high and this evidently enhanced the effect of both contrast in this study and noise in the previous study, so that the effects of other variables were literally "washed out" or obscured by the high light level. There is also a significant effect of transfer function for this group. The blurring of edges, again an effect similar to that of noise and contrast, led to higher complexity judgements.

The analysis for Group 3 is presented in Table 9. The only significant effect here is the interaction between contrast and shading. An examination of the data shows that under conditions of normal contrast, the addition of shading raises complexity ratings but when contrast is lowered, this effect is reversed or disappears probably because the shading is less visible in the low contrast conditions. The significant third order interaction that was observed for Group 1 was not found for Group 3, although the conditions under which the data was collected were comparable.

TABLE 7

ANALYSIS OF VARIANCE FOR STUDY II, GROUP 1

SOURCE	SUM OF SQUARES	DF	MEAN SQUARE	F-RATIO	PROB
Betw Err	304.87	20	15.24		
Contrast (A)	15.02	1	15.02		
W/N Err 1	303.36	20	15.17		
TRN FNC (B)	1.24	1	1.24		
W/N Err 2	45.30	20	2.26		
SHADING (C)	.87	1	.87	3.50	
W/N Err 3	5.00	20	.25		
Edges (D)	4.82	2	2.41		
W/N Err 4	179.09	40	4.48		
AB	.24	1	.24		
W/N Err 5	16.13	20	.81		
AC	.01	1	.02		
W/N Err 6	10.69	20	.53		
AD	1.33	2	.67	1.09	
W/N Err 7	24.41	40	.61		
BC	.02	1	.02		
W/N Err 8	11.19	20	.56		
BD	1.06	2	.53	1.18	
W/N Err 9	18.02	40	.45		
CD	1.00	2	.50	1.16	
W/N Err 10	17.25	40	.43		
ABC	.72	1	.72	1.26	
W/N Err 11	11.32	20	.57		
ABD	.30	2	.15		
W/N Err 12	20.95	40	.52		
ACD	2.33	2	1.17	1.86	
W/N Err 13	25.08	40	.63		
BCD	.76	2	.38		
W/N Err 14	20.65	40	.52		
ABCD	1.59	2	.79	3.97	$<.02$
W/N Err 15	8.00	40	.20		
TOTAL	1052.66	503	2.09		

TABLE 8

ANALYSIS OF VARIANCE FOR STUDY II, GROUP 2

SOURCE	SUM OF SQUARES	DF	MEAN SQUARE	F-RATIO	PROB
Betw Err	240.77	15	16.05		
Contrast (A)	14.65	1	14.65	6.18	<.02
W/N Err 1	35.56	15	2.37		
TRN FNC (B)	21.56	1	21.56	7.50	<.01
W/N Err 2	43.14	15	2.88		
Shading (C)	.32	1	.32	1.12	n.s.
W/N Err 3	4.23	15	.28		
Edges (D)	.27	2	.14		
W/N Err 4	62.65	30	2.09		
AB	.01	1	.003		
W/N Err 5	5.87	15	.39		
AC	.06	1	.06		
W/N Err 6	4.31	15	.29		
AD	.81	2	.41		
W/N Err 7	17.10	30	.57		
BC	.44	1	.44		
W/N Err 8	7.10	15	.47		
BD	.14	2	.07		
W/N Err 9	8.77	30	.29		
CD	.89	2	.45	2.01	n.s.
W/N Err 10	6.69	30	.22		
ABC	.21	1	.21		
W/N Err 11	5.16	15	.34		
ABD	.27	2	.14		
W/N Err 12	25.98	30	.86		
ACD	.14	2	.07		
W/N Err 13	8.10	30	.27		
BCD	.58	2	.29		
W/N Err 14	9.00	30	.30		
ABCD	.75	2	.38		
W/N Err 15	12.50	30	.42		
TOTAL	538.06	383	1.40		

TABLE 9

ANALYSIS OF VARIANCE FOR STUDY II, GROUP 3

SOURCE	SUM OF SQUARES	DF	MEAN SQUARE	F-RATIO	PROB
Betw Err	35.15	9	3.90		
Contrast (A)	6.02	1	6.02		
W/N Err 1	58.90	9	6.54		
TRN FNC (B)	6.02	1	6.02		
W/N Err 2	101.90	9	11.32		
Shading (C)	.02	1	.02		
W/N Err 3	3.40	9	.38		
Edges (D)	5.61	2	2.80		
W/N Err 4	76.72	18	4.26		
AB	.00	1	.00		
W/N Err 5	5.08	9	.56		
AC	2.40	1	2.40	6.78	.03
W/N Err 6	3.18	9	.35		
AD	.66	2	.33		
W/N Err 7	9.18	18	.51		
BC	.07	1	.07		
W/N Err 8	2.18	9	.24		
BD	.11	2	.05		
W/N Err 9	8.72	18	.48		
CD	1.01	2	.50	1.33	
W/N Err 10	6.82	18	.38		
ABC	.42	1	.42	1.41	
W/N Err 11	2.67	9	.30		
ABD	.32	2	.16		
W/N Err 12	10.34	18	.57		
ACD	2.32	2	1.16	2.51	
W/N Err 13	8.34	18	.46		
BCD	1.11	2	.55	1.26	
W/N Err 14	7.89	18	.44		
ABCD	1.26	2	.63	1.64	
W/N Err 15	6.91	18	.38		
TOTAL	374.73	239	1.57		

It is clear from a consideration of the results from the three groups that, although there were a few significant effects, the variables manipulated in Study II had very little effect on the judgement of complexity. The only variable that seems worthy of consideration in further research might be that of contrast but its effect is similar to that of noise, and they probably can be considered equivalent.

Groups 1 and 2 were asked to give brief descriptions of complex and simple scenes. Among those who mentioned noise as being an important characteristic, 13 (11 in Group 1 and 2 in Group 2) said it made a scene more complex and 8 (all from Group 1) said it made a scene less complex. This was one of the reasons for changing the instructions for Group 3. Contrast was also mentioned by 14 people as making a scene more complex and 11 people said it made a scene less complex. 13 people mentioned external edges and 8 mentioned internal edges. A number of people mentioned that the clarity of a scene had a lot to do with their complexity judgements, although some felt it made a scene less complex while others felt it made a scene more complex.

Discussion and Conclusions

The results indicate that it is possible to make meaningful complexity judgements based on CIG simulated scenes, since the parameters that were varied had significant effects. However, there are some differences between complexity judgements of the CIG scenes obtained in the present study and the judgements obtained from either laboratory-generated drawings or real-world stimuli. In the first place, the range of responses is somewhat smaller for the stimuli in the present

study. This can be seen most clearly in a comparison of the means and standard deviations in Table 2 with the mean complexity ratings from de Groot's (1978) data presented in Table 1. The ranges in the former are 2.9 and 1.4 while in the latter the ranges are 5.8 and 5.1. The higher ranges are more commonly reported in the literature (Attneave, 1957; Berlyne, 1974: LeMay and Aronow, 1977), so that the results for the present study are unusual in this respect. This is largely due to the fact that only one scene was simulated. This leaves out the major source of variability found by de Groot (1978), namely different numbers and sizes of man-made objects. While it is not surprising that limiting the sample of stimuli to variations on a single scene should limit variability in the responses, it should be borne in mind that this was done deliberately in order to discern the effects of those variables which could be controlled in the simulation. The introduction of more than one scene would have affected complexity judgements without adding any new information about computer simulation and would probably have masked the effects of edges in the same way that noise did for the first two groups. If anything, the effects of computer simulated variables such as edges, noise, shading, etc., are probably exaggerated by the procedures used here. It may be concluded, then, that once a basic scene is modeled, further refinements in program complexity make very little difference in perceived complexity. It might be pointed out in this connection that the basic scene used only 222 edges, and could not have been done with fewer edges without leaving out either some scene objects or linear perspective.

Another general comparison which can be made between the data obtained here and previous work is in terms of the reliability of the judgements. De Groot (1978) and LeMay and Aronow (1977, unpublished data)

found very low standard deviations, usually less than 1.0, for small (n=9) and large (n=50) groups of subjects. Table 2 shows relatively high standard deviations for the present study indicating less agreement among individuals as to complexity level. This is not too surprising in view of the close similarity of the scenes, but it does seem to indicate that subjects had some difficulty making judgements. The greater spread among individuals, however, was not enough to prevent significant effects for noise and number of internal edges which were strong enough to overcome the increased variability.

It is clear from the results of this study that noise in the display operates to raise the level of judged complexity. This is true even for Group 3 where the subjects have, in effect, been instructed to ignore the noise. In the sense that noise masks some of the detail that would result in higher complexity judgements, it might be expected to have the opposite effect, especially on scenes with many internal edges, and there is some evidence for this in the interaction of noise with internal edges for Group 1. However, most of the subjects seem to consider noise an integral part of the scene and perceive it as adding to scene complexity. In debriefing disucssions, many subjects mentioned that noise added to the difficulty of discerning the detail, thus making the scene more complex. In this sense, they seemed to be looking at the slides as part of a recognition task, even though the instructions asked only for a complexity rating. The recognition of targets will, of course, be a major interest in future research with CIG simulation, and is therefore of considerable interest here. In fact, an important outcome of this study is the scaling of complexity in CIG displays for use in subsequent study to determine the degree of complexity necessary for the basic tasks of detection, recognition, and identification.

If noise increases judged complexity and also increases the difficulty of detecting detail, it might be supposed that more complex stimuli would result in less adequate performance on the tasks mentioned. The literature on this topic has been summarized earlier, and indicates that different results may be expected for different tasks. These results probably depend to a considerable extent on the type of stimuli used and, in particular, on whether the complexity is seen as part of the figure or of the ground. Subjects in this study evidently saw complexity as a unitary characteristic, and therefore added the noise to the rest of the scene. It might be necessary in future research to separate figure and ground, and define complexity separately for each.

The general finding that an increase in the number of edges results in an increase in complexity ratings is in line with the results of previous research (Attneave, 1957; Chipman, 1972; LeMay and Locher, 1978), although it is somewhat surprising that only the effect of internal edges, and not that of external edges, was significant. In studies with random polygons, external edges are frequently the only ones varied, and are the main determiners of judged complexity. In the present study, however, the lack of significance of external edges might be due to the small size of the increment relative to the basic stimulus configuration. As mentioned above, the basic scene was modeled using 222 edges, and adding 59 and 151 edges respectively to make the next two levels of external edges. In other words, the scene is already complex enough so that adding more edges does not add much to perceived complexity. This is borne out by subjects' comments during debriefing. Many of them attributed the addition of edges to a change in style of architecture not related to complexity. The situation with internal edges, however, is quite different. The first level here

has no internal edges at all so that the addition of 59 edges makes for a significant difference in judged complexity. Even here, however, the further addition of another 92 edges does not make for an equal difference in complexity rating, i.e., there seems to be a non-linear trend in the data such that increasing numbers of edges yeilds decreasing returns in complexity ratings. There are not enough data points to test this here, but it might be pursued in the future.

It should be pointed out again, however, that the number of external edges was minimal for scene generation, and could not be further reduced without destroying the scene, and that this is an unavoidable characteristic of scene generation. It simply takes more edges to outline the objects in a landscape than to fill in their surface details. The practical significance of the present result is that, once a basic scene has been modeled with as few edges as possible, additional edges should be used to add surface details such as doors and windows, rather than to enhance the outlines of cultural objects. This recommendation will have to be extended to the detection, recognition, and identification tasks but it is expected that internal edges will prove the more important of the two incremental variables.

The results concerning the effects of the addition of a transfer function, the reduction in contrast, and the addition of shading to the display are extremely difficult to interpret in the light of the appearance and disappearance of significant effects in each of the three groups. Perhaps the most parsimonious explanation is that the effects are examples of Type I errors, and that none of these variables is worth including in future research or in practical use of CIG simulation. At least none of these variables seems to influence complexity judgements very much, although their effect on other tasks, especially identification, should probably not be dismissed so early.

One possible explanation for the lack of significant results is that the data for Study II was always collected after that for Study I, rather than alternating their order of presentation for different groups. This was done because previous research had indicated the overriding importance of edges and noise as determinants of complexity, and this data was considered too important to risk interference by fatigue or boredom. Some subjects, in fact, did indicate that they had used different criteria for the second set of stimuli. This might be overcome in the future by using different groups of subjects, or simply repeating the instructions to the subjects between each set of stimuli. Previous experience with data collection of this sort, however, indicates that very little change could be expected from this alteration in procedure. It is therefore probably safe to conclude that contrast, transfer function (at least of the kind used here), and shading have little or no effect on perceived complexity.

References

Attneave, F. Some informational aspects of visual perception. Psychological Review, 1954, 61, 183-193.

Attneave, F., and Arnoult, M.D. The quantitative study of shape and pattern perception. Psychological Bulletin, 1956, 53, 452-471.

Attneave, F. Physical determinants of the judged complexity of shapes. Journal of Experimental Psychology, 1957, 53, 221-227.

Berlyne, D.E. The influence of complexity and change in visual figures on orienting responses. Journal of Experimental Psychology, 1958, 55, 289-296.

Berlyne, D.E. Conflict, Arousal, and Curiosity. New York: McGraw-Hill, 1960.

Berlyne, D.E., and Parham, L.C.C. Determinants of subjective novelty. Perception and Psychophysics, 1968, 3, 415-423.

Berlyne, D.E. Studies in the New Experimental Aesthetics. New York: John Wiley & Sons, 1974.

Bunker, M., and Heeschen, R. Airborne Electro-Optical Sensor Simulation, AFHRL-TR-75-35, 1975 (July).

Bunker, M., and Heartz, R. Perspective Display Simulation of Terrain. AFHRL-TR-76-39, 1976 (June).

Bunker, M., and Ferris, N. Computer Image Generation Imagery Improvement: Circles, Contours, and Texture. AFHRL-TR-77-66, 1977 (September).

Chipman, S.F. The perceived complexity of visual patterns. Unpublished doctoral dissertation, Harvard University, 1972.

Chipman, S.F., and Mendelson, M. The development of sensitivity to visual structure. Journal of Experimental Child Psychology, 1975, 20, 411-429.

De Groot, S. Human Factors Aspects of Low Light Level Television and Forward Looking Infrared Sensor Displays. AFOSR-77-3242, 1978 (January).

Gibson, E.J., Gibson, J.J., Pick, A.D., and Osser, H. A developmental study of the discrimination of letter-like forms. Journal of Comparative and Physiological Psychology, 1962, 55, 897-906.

Glanzer, M., and Clark, W.H. The verbal-loop hypothesis: Conventional figures. American Journal of Psychology, 1964, 77, 621-626.

Goedert, E.R., Rodwan, A.S., and Salmon, P.G. Specific and general contour adaptation effects found using a selective adaptation paradigm. Perceptual and Motor Skills, 1977, 44, 551-562.

Homa, D., and Coury, B. Perceptual selectivity and the fate of unemphasized information in a stimulus complex. Memory and Cognition, 1977, 5, 347-354.

Kirk, R.E. Experimental Design: Procedures for the Behavioral Sciences. Belmont, Calif.: Brooks/Cole, 1968.

Leckart, B., and Bakan, P. Complexity judgements of photographs and looking time. Perceptual and Motor Skills, 1965, 21, 16-18.

LeMay, M. The perception of patterns of varying complexity in two and three dimensions. Dissertation Abstracts, 1972, 32, 5487-B.

LeMay, M., and Aronow, E. Some determinants of lingering behavior in a suburban shopping center, Perceptual and Motor Skills, 1977, 45, 3.

LeMay, M., and Locher, P. Some determinants of the judged complexity of houses. Paper presented at the meeting of the Easter Psychological Association, Washington, April, 1978.

Meyer, B. Exploratory analysis of visual scale, complexity, and proportion factors in architectural form perception. Dissertation Abstracts International, 1973, 34 (3-B), 1260.

Nygard, J.E., Slocum, G.K., Thomas, J.J., Skeen, J.R., and Woodhull, J.G. The Measurement of Stimulus Complexity in High-Resolution Sensor Imagery. AMRL-TDR-64-29, 1964 (May).

Pasnak, R. A quantitative method for predicting the discriminability of human faces. Paper presented at the meeting of the Eastern Psychological Association, Philadelphia, April, 1969.

Pasnak, R. Pattern complexity and response to distinctive features. American Journal of Psychology, 1971, 84, 235-245.

Saflofske, D.H. Aesthetic complexity and exploratory behavior. Perceptual and Motor Skills, 1975, 41, 363-368.

Seman, W.A., Pasnak, R., and Tyer, Z. Discriminability and identifiability of the human hand based on its quantified complexity. Perceptual and Motor Skills, 1976, 42, 287-293.

Sternberg, J.J., and Banks, J.H. Search Effectiveness with Passive Night Vision Devices, June, 1970 B ESRL-TR-1163, AD714 206, Army Behavior and Systems Research Laboratory, Arlington, VA.

Tone, T., and Muraoka. The relationship between length of verbalization and complexity of figures: The verbal-loop hypothesis. Japanese Psychological Research, 1975, 17, 213-215.

Uttal, W.R., and Tucker, T.E. Complexity effects in form detection. Vision Research, 1977, 17, 359-365.

Vanderplas, J.M., and Garvin, E. The association value of random shapes. Journal of Experimental Psychology, 1959, 57, 147-154 (a).

Vanderplas, J.M., and Garvin, E. Complexity, association value, and practice as factors in shape recognition following paired-associates training. Journal of Experimental Psychology, 1959, 57, 155-163 (b).

Wilson, W.H. Visual selection: the effects of number of encodable elements and physical complexity. Perception and Psychophysics, 1976, 20, 205-209.

A MODEL OF PERFORMANCE EFFECTIVENESS IN THE AIR FORCE MAINTENANCE SYSTEM

DESIGN REPORT

by

Hewitt H. Young

USAF/ASEE/OSU SUMMER FACULTY FELLOW

August 1978

ADVANCED SYSTEMS DIVISION
AIR FORCE HUMAN RESOURCES LABORATORY
WRIGHT-PATTERSON AIR FORCE BASE, OHIO 45433

A MODEL OF PERFORMANCE EFFECTIVENESS
IN THE AIR FORCE MAINTENANCE SYSTEM:

Design Report

Hewitt H. Young, Professor of Engineering
Arizona State University, Tempe, Arizona
USAF/ASEE/OSU Summer Faculty Fellow 1978

Abstract

A major 1970's objective of the U.S. Department of Defense is to improve the effectiveness of weapon system maintenance and maintainability. To meet this objective, the U.S. Air Force, as well as the other military services, is expending considerable resources in determining the content of maintenance tasks, in determining the impact of equipment design decisions upon these tasks, in streamlining selection and training procedures for maintenance personnel, and in providing well-documented technical information on equipment.

This report considers the design of a maintenance performance effectiveness prediction and evaluation model. Development of such a model would be useful for: (1) inclusion in life-cycle-cost predictive models, (2) AF Wing and Squadron management of maintenance activity, (3) AF Command evaluations of Wing and Base maintenance performance, and (4) determination of future concentration areas for human resources efforts relating to weapon system maintenance. The purposes of the project herein reported on were: (1) to isolate and identify the primary factors which affect maintenance performance effectiveness, (2) to describe predictive models of maintenance manpower effectiveness from past research, and (3) to establish an experimental design for collecting and interpreting data needed for a generalized model. The study was oriented towards maintenance technician performance and human resource factors.

A separate Preliminary Report provides a literature review and summary of past Air Force research relating to maintenance performance. In this report several predictive models from past research are further analyzed and combined, leading to a taxonomy of factors which are indicative of maintenance performance effectiveness and to a generalized model formulation.

Thirty-six condition variables are suggested as predictors of technician squadron performance, the latter measured by averaged speed of performance and averaged accuracy (quality) of performance. The condition variables are classified in the Taxonomy as (1) Equipment Reliability/Maintainability Factors, (2) Maintenance Equipment and Technical Information Factors, (3) Technician Experience, Skill and Knowledge Factors, and (4) Technician Productivity/Morale Factors.

An experimental design for an AF Command study of the variable relationships is established. The purpose of such experiment would be to establish the proper weights or equation coefficients across the 36 predictor variables in the model. It is hypothesized that different weights would be appropriate for different end-item equipments, different squadrons (organizational, field, avionics, munitions, depot), and different Commands. A second purpose of the study would be to establish normative values for both the predictor variables and the performance measures, and thus further prioritize the relative importance of the 36 predictor variables. Some of these variables may prove insignificant or of slight significance and unworthy of use in the model.

The experimental design incorporates a statistical data collection and analysis methodology, using inputs in the forms of survey instruments, supervisory ratings of technician performance and climatological data. In the proposed instrument for technician input, items are extracted from proven source instruments such as the Gordon Personal Profile, the AF Occupational Attitude Inventory, the Organizational Climate Inventory, and the Goldman Group Cohesiveness/Morale Survey.

TABLE OF CONTENTS

LIST OF TABLES

GLOSSARY OF TERMS

316XO TECHNICIAN - Missile Electronics Specialist AFSC

426XO TECHNICIAN - Munitions Weapons Mechanic AFSC

463XO TECHNICIAN - Nuclear Weapons Specialist AFSC

ASCENDANCY TRAIT - Assertive in relationship with others; desire to assume an active role.

AFSC - Air Force Specialty Code

ASSIGNMENT LOCALITY - Location, climate, community atmosphere

EMOTIONAL STABILITY TRAIT - Well balanced and relatively free from anxieties and nervous tensions

EXTRINSIC INCENTIVE - External reward which provides job satisfaction, such as pay

FATIGUE TRAIT - Subjectiveness to feelings of weariness

FATIGUE STATE - Current feelings of weariness

GROUP HOMOGENEITY OF ATTITUDE - Common purpose and goals

INTRINSIC INCENTIVE - Internal reward which provides job satisfaction, such as fulfillment

HIGH PERFORMER - Upper 50 percent of subjects on a selected measure of performance

JOB CURIOSITY TRAIT - Interest in discovering and learning all facets of a job

LBDQ - Leadership Behavior Description Questionnaire

LEADERSHIP - Styles, patterns of supervisor behavior

LOW PERFORMER - Lower 50 percent of subjects on a selected measure of performance

MSEP - Maintenance Standardization and Evaluation Program

MSET - Maintenance Standardization and Evaluation Team; MSET tasks are those subject to Command Team evaluation

ORGANIZATION CLIMATE/IDENTITY - Atmosphere which creates a feeling of belonging as a valuable member of a working team

ORGANIZATIONAL CLIMATE/REWARD - Atmosphere which creates a feeling of being rewarded for job's well done

ORGANIZATIONAL CLIMATE/RISK - Atmosphere which creates a feeling of riskiness or uncertainty about job and/or organization, including job safety hazards

ORGANIZATIONAL CLIMATE/STANDARDS - Atmosphere which creates a feeling for the perceived importance of implicit and explicit goals and performance standards

ORGANIZATIONAL CLIMATE/STRUCTURE - Atmosphere which creates a feeling of group constraints via rules, regulations, red tape

ORGANIZATIONAL CLIMATE/WARMTH - Atmosphere which creates a feeling of general good fellowship

ORGANIZATIONAL IDENTIFICATION - Feelings of association with and support from the organization

PAY AND BENEFITS SATISFACTION - Feelings of technician about these extrinsic rewards

PERSISTENCE TRAIT - Sees work through to completion

PROFESSIONAL IDENTIFICATION - Good feelings about job speciality as important and necessitating special skills

RANK - Position level in AF organizational structure

RESPONSIBILITY TRAIT - Positive and perservering in using own judgements; determined; reliable

SAINT - Systems Analysis of Integrated Networks of Tasks; a computerized simulation methodology

SAT - SATisfactory rating in quality of work inspection

SOCIAL STATUS - Feelings about the importance of occupation within society; a perceived level of society acceptance and status

SRAM - Short Range Attack Missile

WUC - Work Unit Code for AF equipment system, subsystem, or component

INTRODUCTION

A major 1970's objective of the U.S. Department of Defense is to improve the effectiveness of weapon system maintenance and maintainability. To meet this objective the U.S. Air Force, as well as the other Services, is expending considerable resources in defining the content of maintenance tasks, in determining the impact of equipment design decisions upon these tasks, and in streamlining the selection and training procedures for maintenance personnel.

The official maintenance organization and structure are defined by Air Force Manual 66-1. There are three organizational levels known as Organizational, Intermediate, and Depot. Most maintenance activity can be classified as either electronic or mechanical, and there are seven major types of maintenance activity (Foley, 1976).

The Air Force Maintenance System has three primary objectives: (1) to keep AF equipment operational a maximum percent of any given time period, (2) to perform maintenance activity in the most effective manner in terms of time, cost, and quality criteria, and (3) to contribute to future improvement in the maintainability design of AF equipment. Since these objectives are not exclusive, the critical balance among these objectives is a significant management problem.

Life-cycle-cost models for new weapon system designs include operational and maintenance categories. The Grumman and Lockheed Aircraft Corporations have developed LCC models of considerable promise. However, the predictor variables for maintenance activity are mostly equipment-oriented rather than man-oriented. For instance, in the Grumman model the life-cycle-cost factor for base level maintenance on aircraft landing

gear is a function of the Take-Off Gross Weight-Maximum and the Rate of Aircraft Descent at Landing. Such equations predict the amount of maintenance cost that may be occurred over the life-cycle of the equipment, based on design and operational parameters. However, life-cycle-cost of a weapon system also depends heavily upon the performance level of the airmen and civilians who maintain the system, as well as upon the maintainability design of the system. Yet predictive models of manpower performance effectiveness in the Air Force exist only in a few isolated research studies and for particular subsystems, tasks, and job codes. Development of a generalized Air Force predictive model of maintenance manpower effectiveness would be useful for:

(1) inclusion in life-cycle-cost models;

(2) wing and squadron management of maintenance activity;

(3) Command evaluation and improvement of base and wing maintenance performance, and

(4) determination of future concentration areas for human resources research relating to weapon system maintenance.

PURPOSES AND SCOPE OF STUDY

The purposes of this study are therefore: (1) to isolate and identify the primary factors which affect maintenance performance effectiveness, (2) to describe predictive models of maintenance manpower effectiveness from past research, and (3) to establish an experi-

mental design for the collection and interpretation of data needed for a generalized predictive model. This study is essentially a proposal for an expanded study, to be conducted by the Human Resources Laboratory, WPAFB, in conjunction with one or more Air Force Commands.

BACKGROUND FOR STUDY

Reference is made to the Preliminary Report, which provides an extensive review of Air Force projects and other projects that have been concerned with maintenance performance. The Preliminary Report also summarizes comments received from interviews with a small group of maintenance officers. The end product of the Preliminary Report is a listing of factors which have been shown to affect maintenance performance in selected research studies.

Some additional background material has been obtained since publication of the Preliminary Report, and items pertinent to this study are cited in the following sections. A number of the source materials were obtained through computerized literature searches of the Defense Documentation System (DDS) and the Defense Logistics Studies Information Exchange (DELSIE), secured by the Battelle Research Institute/Columbus. Battelle Institute is under contract with AFHRL/WPAFB for a study of problem areas which affect the proficiency and satisfaction of AF maintenance technicians, with the intent of offering solutions.

One report (LaRue and Metzinger, 1974) describes a comprehensive study of factors which motivate maintenance personnel to work particular shifts. The statistically significant factors proved to be (1) Age, (2) Grade Classification and (3) Salary. The study was primarily concerned with civilian maintenance personnel.

The historical and present roles of the Maintenance Crew Chief were reviewed in another study (Beu and Nichols, 1977). It was noted that AFM 66-1, which emphasizes a decentralized maintenance activity and a strong centralized control system, has diminished the role of the Crew Chief (and organizational maintenance) to launching, recovery, inspection, towing, washing, refueling, oil sampling, and minor replacement activities (tires, lights, panel screws). All major "on-board" maintenance activity necessitates called-in crews from the Field Maintenance, Avionics Maintenance, or Munitions Maintenance Squadrons. The study suggests that dedication and professionalism of the Crew Chief has been lost and that this impacts on the productivity of the base maintenance activity. A conclusion of the study is that more "on-board" maintenance activity should be assigned to organizational maintenance and the Crew Chief. The study includes a description of maintenance airmen morale problems generated within PACAF/IGL and reproduced below:

> When a young man enters the Air Force and finds he is qualified for the aircraft maintenance career field, he is filled with a sense of pride to know that he is going to work on aircraft; after all, that's what the Air Force is all about: aircraft and flying. He undergoes training and receives the finest technical instruction in the world and, upon graduation from school, he is assigned to an aircraft maintenance organization.

He then meets his flight chief. Now, eager to put the training he has received to use, he faces the realities of aircraft maintenance: the long, miserable duty day (routinely 12 hours long), missed meals, never-ending flow of work, the uncertainty of planning for a weekend off, adverse weather conditions, no praise or recognition, and constant harrassment from higher headquarters, MSET, ORI teams, and staff assistance visits. He quickly finds that, routinely, the aircraft maintenance career field is critically undermanned at the working level. To his further disillusionment, he becomes aware that aircraft mechanics are among the highest users of drugs and alcohol in the Air Force, and that they have become, for the most part, complacent about their trade.

What can we do to reverse this trend? We certainly don't need more catchy little aphorisms such as "doing more with less," since that is exactly what we have been doing in aircraft maintenance since 1947. Steps must be taken to either reduce the workload or increase the number of qualified mechanics available to perform the quality maintenance we must have; and, a firm duty period - perhaps three 8-hour shifts, or nine shifts in eight days followed by a four-day

break. What we most assuredly need is a realistic look at working conditions of all aircraft mechanics...with a view toward building pride of workmanship, diligence, and dedication in aircraft maintenance from which we will all reap the benefits.

(Beu & Nichols, 1977)

Other researchers (Carpenter-Huffman and Rostker, 1976) have studied the relevance of training for the maintenance of advanced avionics. They conclude that, since a majority of AF maintenance technicians serve only the enlistment term, specific job performance and only job performance should be taught; that job performance should be taught formally and at an operating base and when needed. They also conclude that initial formal training should be alternated with job experience. The authors provide a taxonomy of specialty shop content described as follows:

Field Maintenance - propulsion, pneudraulics, fuel, electrics, environmental control, welding, sheet metal, AGE, corrosion control

Avionics Maintenance - communications, navigation, radar, computer, electrics

Munitions Maintenance - conventional weapons, special weapons, forward firing, weapons release, storage area

A study of job characteristics and job attitudes at a TAC fighter aircraft maintenance complex (Guthrie 1977) showed that maintenance airmen have a lower degree of job satisfaction than for other AFSC's, citing such factors as work environment, growth potential, challenge, etc. As survey tools Guthrie used the Hoppack Job Satisfaction Measure and the Hackman and Oldham Job Diagnostic Survey.

One of the most interesting studies was conducted by Perceptronics Corporation (Drake, Sanders, Crooks, and Weltman, 1977). They conducted a comparative analysis of organizational factors affecting military maintenance. Conclusions were stated as, "Results of our analysis indicate that the biggest payoff in improving military maintenance effectiveness and efficiency is not in introducing additional incentives, but rather in reducing or eliminating the existing disincentives." The authors recommend that: (1) job enrichment activities be instituted to modify the mechanic's job, and (2) an effort be made to reduce the impact of necessary interruptions on maintenance activities.

The performance and satisfaction of Air Force personnel are also affected by the types of feedback provided to them (Pritchard and Montagno, 1978). The authors note that "work motivation" has received far less attention than "selection and training," yet is of equal or greater importance to airman job performance and satisfaction. They further note that the non-monetary intrinsic rewards need more AF attention, and that proper feedback can be an important intrinsic award. Dimensions of feedback are stated as (1) positive vs negative, (2) timing, (3) specificity, (4) evaluative-nonevaluative, and others given in the report. For high motivation, the authors suggest from their research that <u>all</u> of the following conditions must be present:

1. The person must feel that he can influence his performance by his effort

2. He must see that high performance leads to high rewards

3. He must value the rewards

Another recent study sought the factors which affect performance ratings by supervisors (Wiley and Hahn, 1977). A large volume of data was collected on sources of training on tasks, on the motivational aspects of doing tasks, and on retention of skills. The AFSC's studied included Telecommunications Operations Specialists (AFSC 29150), Ground Radio Communications Equipment Repairmen (AFSC 30354), and Aircraft Maintenance Specialists-Jet (AFSC 43151C). The basic findings of the study were that:

1. Raters can agree on task performance evaluations to a statistically significant degree.

2. Raters agree better on rating overall job performance than on rating the performance of single tasks.

3. Incumbents are poorer sources of task ratings than peers or supervisors.

4. Peers can be substituted for supervisors as performance evaluators without great loss in reliability.

5. A few task ratings taken together account for a substantial percentage of the overall performance rating variance.

6. Either, or both, aptitude data and demographic data (such as grade and length of service) account for much less of the overall evaluation of an airman's performance than can be had from ratings on as few as five tasks.

7. It is likely that some of the unaccounted-for overall performance variance was attributable to the attitudes and interests of the incumbents. (Only preliminary analyses could be made of these data, but the prediction results were statistically significant.)

8. It is possible that small improvements can be made in selecting personnel for certain task assignments by means of tests not in current use; but these improvements will be slight, at best.

9. Marked differences distinguished the 304X4 AFSC performance ratings from the other two specialties; they were internally more reliable and better able to predict overall performance ratings.

10. Use of the top of the rating scale was frequent when performance ratings were made on easier tasks, and on performance in AFSCs with lower aptitude requirements.

11. Since measurability was better for more difficult tasks, with less use of the top of the rating scale, the AFSC with high aptitude incumbents received the lowest mean performance ratings of the three specialties.

A TAXONOMY OF MAINTENANCE MANPOWER EFFECTIVENESS

Based on an extensive literature review, interviews, and observations in the 4950th Test Wing at WPAFB, interviews at AFIT/LGSQ with officers having extensive experience in maintenance, discussions with Battelle/ Columbus personnel, and discussions with AFHRL/ASR staff, the writer has prepared a tentative Taxonomy of factors which impact upon maintenance manpower effectiveness. Such manpower effectiveness is influenced by (1) Equipment Reliability/Maintainability Factors, (2) Maintenance Equipment (test, handling, tools, etc.) and Technical Information Factors, (3) Technician Experience, Skill and Knowledge Factors, and (4) Technician Productivity/Morale Factors. The Taxonomy is shown in Table 1.

The factors listed in the Taxonomy have either (1) been statistically tested for particular equipment work unit codes (WUC's) and AFSC's and found to significantly contribute to task performance measures of speed and/or accuracy, or (2) been ranked as important contributors to performance effectiveness by airmen and/or their supervisors in non-statistical studies. Performance effectiveness is herein defined as a function of the quantity and quality of maintenance activity. The factors listed in the Taxonomy are intended to cover all levels of maintenance (organizational, intermediate, depot) and all types of Air Force Equipment; however, a few factors may relate primarily to base maintenance (organizational/intermediate) and others to depot maintenance.

Performance Measurement Factors

Performance effectiveness may be considered a composite of performance rate (speed of work) and performance accuracy (quality of work). Both are

TABLE 1

A Taxonomy of Air Force Maintenance Manpower Effectiveness

Equipment and Information Resources

Equipment Reliability/Maintainability Factors	Maintenance Equipment and Technical Information Factors
Maintainability	**Handling**
Weight/size of equipment system	Malfunction occurrence level in handling equipment
Weight/size of subsystem and components	Weight/size of end item equipment
Access for test and check	Location of equipment at flightline
Clearance for removal/replacement	**Testing and Repair**
Number and arrangement of internal components	Design efficiency of test equipment
Number, location, and arrangement of test points	Level of information received from tests
Reliability	Type and amount of test equipment information provided
Design operating parameters of equipment	Length of procedural sequence for test
Average flight or operating hours between servicing of equipment	Logic information for diagnosis
Level of discrepancy reporting by aircrew (aircraft only)	Ease of hand tool use
Spares availability	Adequacy and completeness of technical information on equipment

Table 1 (Cont)

A Taxonomy of Air Force Maintenance

Manpower Effectiveness

Human Resources

Technician Experience, Skill, and Knowledge Factors

Experience
AFSC level (3,5,7, etc.)
Rank
Months in career field

Skills (supervisor/observer ratings)
For components worked on
For test equipment used
For use of test equipment in general
For equipment repair

Knowledge (supervisor/observer ratings)
Of maintenance procedures
Of equipment handling procedures
Of use of equipment when operational
Of equipment maintenance procedures in general

Technician Productivity and Morale Factors

Organizational Climate and Group Morale
Competency of supervision
Supervisory conditioning of tasks
Structure / Warmth / Standards / Identity / Risk
Group satisfaction of individual motives
Satisfaction with interpersonal relationships
Team cohesiveness

Personal Traits and Motivators
Job curiosity trait
Persistence trait
Emotional stability trait
Fatigue trait
Organizational identification
Professional identification
Responsibility trait
Self-starter trait
Ascendency trait

Operational and Environmental Conditioners
Pay and benefits, as perceived
Assignment locality and climate
Airmen/civilian relationships
Participation in interest/service clubs
Social status of occupation, as perceived
Lighting/noise/workplace size/ clothing

important, with the latter taking precedence when sortie availability and operational safety are affected. An associated measure is job satisfaction.

Performance rate measures include (1) the inverse of average times to perform selected maintenance tasks, and (2) supervisor or observer ratings of average performance speeds on selected maintenance tasks. The first is an objective measure, available through summary reports from the AFTO 349 task reporting for or by time-study observation. The second is a subjective measure which necessitates special supervisory input. The latter technique was employed in the recent study on nuclear missile maintenance handling tasks (Sauer, Campbell, Potter, and Askren, 1977). Performance rates can also be predicted from a computer simulation model, as was employed in a study on nuclear systems safety (Askren, Campbell, Seifert, Hall, Johnson, and Sulzen, 1976).

Performance accuracy measures include (1) rated SAT scores from Personnel Evaluations and Technical Inspections conducted under MSEP (Maintenance Standardization and Evaluation Program), reported monthly, and (2) supervisor or observer ratings of performance accuracy on selected maintenance tasks.

Other performance measures have been tried. The following additional measures were used in the study of troubleshooting electronic equipment (Meister, Finley, and Thompson, 1971):

1. Whether T.O. used or not

2. Technician's understanding of problem (observer's rating)

3. Number of unique diagnostic checks made

4. Number of repeated diagnostic checks made

5. Number of components removed/replaced

6. Observer's rating of technician efficiency

7. Technician's rating of task difficulty

8. Number of times assistance is required by technician

9. Number of components actually worked on

10. Maintenance diagnostic strategy

Although relationships were found between the above performance measures and some of the predictor variables used, this writer believes that all of the above are indicative of performance speed and accuracy. Therefore, only the two performance measures of speed and accuracy are suggested for the proposed study.

Evidence From Past Research Concerning the Driver Variables

Study A. The study on two autopilot systems (Meister, Finley, and Thompson, 1971) involved data collection from the organization and intermediate (avionics) maintenance squadrons at March AFB and WPAFB. Table 2 shows combined results of the statistical analyses conducted on the experimental data for the performance factor of "time to perform maintenance task." Performance time data was collected directly from the AFTO 349

TABLE 2

Summary Statistics from Study of Maintenance Job Performance
on Two Autopilot Systems (Compiled from Meister, Finley, and Thompson, 1971)
Performance Time Predictors (Observation)

Ident.	Predictor Variable	Multiple Regression Coeff. Shop (inter) March AFB (B3.30)	Multiple Regression Coeff. Flightline (B-15.5)WPAFB	Multiple Regression Coeff. Shop (B2.64)	Correlation Coefficients Correl. R's _ .30 March AFB	Correlation Coefficients WPAFB
29	Clearance Remove/Replace				-.47	
37	Type Diag. Information	$\boxed{S}$		* - .070	$\boxed{-.65}$	-.33
47	Assessibility Int. Comp.			* - .623	-.34	-.40
51	Type Test Equip. Info.			* -1.584		-.37
53	Amount Test Equip. Info.				-.35	
55	Design of Test Equip.				-.48	
57	T.O./Checklist Info.	- .557* $\boxed{S}$	7.873	*	$\boxed{-.98}$	-.32
71	AFSC Level			* + .892	.30	.47
73	Rank		4.158	*	.47	.38
75	Experience				-.50	
85	Supv. Rat.-Skill Working on Indiv. Components				.52	-.47
87	Supv. Rat.-Test Equip. Item Used-Skill	$\boxed{S}$			$\boxed{.77}$	-.55
89	Supv. Rat.-Overall Test Equipment Usage	$\boxed{S}$		$\boxed{S}$	$\boxed{.77}$	$\boxed{-.69}$
91	Supv. Rat.-Equip. Repair Performance				.51	$\boxed{-.71}$
93	Supv. Rat.-Procedures	Excluded from regression analyses.			.52	$\boxed{-.72}$
95	Supv. Rat.-Handling Skills				.32	-.48
97	Supv. Rat.-Functions				.58	$\boxed{-.70}$
99	Supv. Rat.-Overall Capability				$\boxed{.61}$	$\boxed{-.68}$

* - Surfaced in Regression Analysis
$\boxed{S}$ - Significant in Anova F Test at ($P < .10$) or better

Selected Predictive Regression Equations [Supervisor ratings excluded]
March AFB
Flight Time - None
Shop - $y = 3.30 - 0.557\,(57) - 0.005(57)^2$
WPAFB
Flightline - $y = -15.55 + 7.873\,(57) - 1.424(57)^2 + 4.158(73) - 0.649(73)^2$
Shop - $y = 2.64 - 0.070(37) - 0.623(47) - 1.584(57) + 0.892(71)$

reporting forms. The correlation coefficients were computed first for the several predictor variables, and those variables with R's $\geq$ .30 were then subjected to further analyses. Data classification analysis on these factors, shown in Table 2, led to a reduction of this list of variables; those left were then included in a multiple regression analysis. The results of the regression analysis and the selected predictive regression equations are also given in Table 2. Supervisory ratings were excluded from the regression equations.

As may be observed from Table 2, six predictor variables (excluding supervisory ratings) surfaced in the multiple regression analyses:

1. The type and amount of technical information provided in TO's, checklists, etc.

2. The rank of the airman

3. The type of test equipment information provided

4. The AFSC level

5. The accessibility of internal components

6. The type of diagnostic (troubleshooting) information provided

In addition, there were high correlations between several of the supervisory ratings and the maintenance task performance time. The study did not include motivational factors, nor was accuracy used as a direct performance measure.

Study B. In the study on missile handling safety (Askren, Campbell, Seifert, Hall, Johnson, and Sulzen, 1976), human resources data and maintenance task data were developed for nuclear system maintenance operations associated with the Short Range Attack Missile (SRAM). The data was then integrated into a computerized network simulation model known as SAINT (Systems Analysis of Integrated Networks of Tasks) to generate predicted maintenance task times and handling task safety hazards. The simulation was shown to provide an opportunity for evaluating the impact of selected predictor variables on the performance measures. In the study, 13 psychologists and human factors engineers participated in scaling the relationships between the performance measures of task time/task safety and the selected predictor variables. The relationships are thus subjective and not based on collected operational data.

The five predictor variables selected for the computer simulation were extracted from 11 human resource factors by a panel of eight human factors specialists. The 11 factors and the 5 selected as most critical to performance time and performance safety are:

Eleven Human Resource Factors	Selected as Most Important Predictors of Performance Time or Safety
1. System Proficiency (Months in SRAM assignment)	*
2. Fatigue	*
3. Environment (Temperature)	*
4. Motivation	*
5. Written Manuals	*
6. Career Field Training and Experience	
7. Organizational Structure	

8. Aptitude

9. Emotional Stability

10. Team Cohesiveness

11. Leadership

Maintenance task data (SRAM Ground Handling Task Times and Task Hazard Scores) were developed subjectively across 2 tasks by interviews with 120 technicians (at three AFB's). The performance time estimates included most likely, minimum and maximum times to complete tasks.

Relationships between the two performance measures and the five predictor variables were then subjectively developed as regression equations by the 13 raters, covering three task areas: transport, checkout, and assemble/disassemble. Thus a total of 30 graphs were created (5 human resource factors X 2 performance variables X 3 tasks).

The SAINT computer simulation model used levels of the human resource factors and the maintenance task data to generate similar performance measures (equations) to those subjectively developed by the 13 raters. A comparison of the simulation data with the estimates of the 13 raters is troublesome as shown in Table 3. Either the subjective estimates of predictor variable effects on performance time were considerably overestimated by the 13 human resources personnel or else the simulation model does not adequately represent the real world. All factors except the ones varied were maintained at baseline levels. In the simulations, the variables were permitted to vary randomly between the following limits based on estimates of the distribution forms:

TABLE 3

Comparison of Performance Time Curves
as Estimated by 13 Human Resources Raters and
as Generated by SAINT Computer Simulation
(Compiled from Askren, Campbell, Seifert, Hall,
Johnson, and Sulzen, 1976)

A. Performance Times vs Number of Fatigue Symptoms for SRAM Missile Handling Task 27 (Load Launcher to Aircraft)

Number of Fatigue Symptoms	Mean Subjective Time Estimates of 13 Raters (Minutes)	Computer Generated Times Using SAINT Simulation
0	54.6	59.9
5 Baseline	63.6	61.1
10	72.6	62.7
15	81.6	
20	90.6	65.7
25	99.6	
30	108.6	68.8

B. Performance Times vs Environment (Temperature) on Three Tasks on SRAM Missile

Temperature Conditions	Task 2 (Transport)[1]		Task 27 (Assemble)[2]		Task 29 (Checkout)[3]	
	Raters	Simulation	Raters	Simulation	Raters	Simulation
-40°F	22.6	15.7	114.0	70.7	483.6	271.3
0°	18.3	15.0	87.6	66.0	358.8	265.5
40°	14.8	14.3	66.0	62.0	275.6	261.4
60° Baseline	14.2	14.2	61.8	61.1	260.0	260.7
90°	15.1	14.4	68.4	62.2	291.2	261.9
120°	19.3	15.2	96.0	67.1	397.8	266.6

1. Task 2 - Transport Payload to IMF, 463XO Crew, Mean Time 140 Minutes

2. Task 27 - Load Launcher to Aircraft, 462XO Crew, Mean Time 60 Minutes

3. Task 29 - Perform Aircraft System Checkout, 316XO Crew, Mean Time 260 Minutes

Predictor Variables	Worst Level	Baseline (Ave.)	Best Level
System Proficiency (X_1)	6	12	30
Quality Written Materials (X_2)	10	80	100
Environment (Temp.) (X_3)	-40	60	120
Number of Fatigue Symptoms (X_4)	30	5	0
Work Motivation Score (X_5)	0	3	6

It is of some interest to combine the subjective estimates of relationships between maintenance performance time and the five human resource factors used in the study. Given the following subjective estimate time equations from the report for Task 29 (Aircraft System Checkout):

Predictor Variable	Subjective Equation	Scale
Proficiency (Months in SRAM)	$y=48.38-4.84X+0.083X_1^2$	$0 \leq X_1 \leq 36$
Written Manuals Quality Index	$y=96.09-1.22X_2$	$0 \leq X_2 \leq 100$
Temperature °F	$y=38.32-1.08X_3+0.0048X_3^2$	$-40 \leq X_3 \leq 120$
No. of Fatigue Systems	$y=-10.96+2.98X_4$	$0 \leq X_4 \leq 30$
Work Motivation Scale	$y=69.90-27.14X_5+1.24X_5^2$	$0 \leq X_5 \leq 6$

where y=% increase or decrease from baseline

A composite equation in time domain, assuming independence and additivity of the equations, would be (baseline time is 260 minutes for Checkout Task):

$$Y \text{ minutes} = 260\left[1+\left[.4834-.0484X_1+.00083X_1^2+.9609-.0122X_2^2+.3832 -.0108X_3+.000048X_3^2-.1096+.0296X^4+.6990-.2714X_5+.0124X_5^2\right]\right]$$

$$= 260\left[3.4173-.0484X_1+.00083X_1^2-.0122X_2-.0108X_3+.000048X_3^2+.0298X_4 -.2714X_5+.0124X_5^2\right]$$

Thus at the baseline levels for all 5 predictor values:

$$Y \text{ minutes} = 260\left[3.4173\overbrace{-.5808+.1195}^{X_1}\overbrace{-.9760}^{X_2}\overbrace{-.6480+.1728}^{X_3}\overbrace{+.1490}^{X_4}\overbrace{-.8142+.1116}^{X_5}\right]$$

$$= 260\left[.9512\right] = 247.312$$

The error between 260 and 247.3 is due to the fact that the predictor variable levels at baseline crossings are not exactly at the integer values selected. It can be seen that the contributions to the predictive equation for maintenance performance time on Task 29, in order of importance, are:

	Contribution	% of Contribution
X_2 - Quality of Written Materials	0.9760	35.4
X_5 - Work Motivation Score	0.7026	25.5
X_3 - Environment (Temperature)	0.4752	17.2
X_1 - System Proficiency (Skills)	0.4613	16.8
X_4 - Number of Fatigue Symptoms	0.1490	5.5

Note that the Quality of Written Materials is twice as important as environmental temperature.

Study C. An extension on Study B above (Sauer, Campbell, Potter, and Askren, 1977) made a more comprehensive analysis of the human resource factors and equipment/environmental factors affecting maintenance performance. First, opinion data from 230 SRAM, Minuteman and Genie maintenance technicians was ranked in order of importance as shown in Table 4. More detail is provided in Appendix A. The rankings reflect on averaged ordering of the opinions of the 230 technicians.

TABLE 4

Ranked Opinion Data from 230 SRAM, Minuteman, and GENIE Technicians on Factors Affecting Maintenance Performance (Extracted from Sauer, Campbell, Potter, and Askren, 1977)

Human Resources Ranks (Rank 1 is high)

Item	Rank
Team Cohesiveness	1
Emotional Stability	2
Fatigue	3
Systems Training/Experience	4
Leadership	5
Motivation	6
Career Field Training	7
Organizational Structure	8
Military Morale/Attitude	9
Aptitude	10
Military Regulations/Procedures	11
Personal Life	12
Respect & Reinforcement from Superiors	13
Communication Between Supervisor & Technicians	14
Feelings of Self-Accomplishment/Pride	15
Maturity of Technician	16
Fellow Crew Members	17
Cooperation Between Shops	18
Shift Work	19

Equipment/Environmental Ranks (Rank 1 is high)

Item	Rank
Equipment Reliability	1
Weather Conditions	2
Operation of Equipment	3
Technical Orders	4
Lighting Conditions	5
Noise Level	6
Equipment Safety Features	7
Work Place Size/Shape	8
Clothing Types	9
Usefulness of Maintenance Equipment	10
Equipment/Parts Availability	11
Equipment Serviceability	12

As was developed in the Preliminary Report, the Sauer, et al. study also employed supervisor rankings of the accuracy and speed of task performance to evaluate an extensive list of possible predictor variables. Table 5 shows the performance speed results of three statistical techniques applied to the data: (1) high performer vs low performer analyses, (2) correlation analyses, and (3) multiple correlation analyses. Table 6 provides similar results for performance accuracy.

Multiplying the regression equation coefficients by the mean values for high performers of the selected predictor variables, the predictor contributors for MSET Task 1101A and 462X0 technicians can be weighed in priority order as follows for ranked speed of performance:

Factor	Coefficient	Mean Value H. P.	Weight
1. Ascendency trait	.2218	20.5	4.5691
2. Group homogeneity of attitude	.2456	13.9	3.4138
3. Assignment locality	.1933	16.0	3.0928
4. Organizational climate/risk	.2241	10.0	2.2410
5. Fatigue trait	.3193	2.3	0.7344
6. Number of extracurricular interest clubs	.1702	0.5	0.0851
7. Number of extracurricular service clubs	.2132	0.24	0.0512

The last three predictor variables have a minor effect on performance speed (or time) when compared with the first four variables.

A similar analysis by accuracy of task performance on MSET 1101A Task (462X0 technicians) leads to the following priority list of predictor variables:

TABLE 5

Statistical Results on Factors Affecting Maintenance Performance Speed Based on Data Collected Across 140 SRAM Technicians AFSC 462XO and 463XO (Compiled from Sauer, Campbell, Potter, and Askren, 1977)

Predictor	Significant High Performer Variables and Multiple Regr. Coefficients			Correlation Coefficients Correl. R's $\geq$ 0.25	
	462XO(B33.1879)	463XO(B70.8817)		462XO	463XO
Years of Service		.2168		.28	(.17)
Months in Career Field				.25	.29
No. of Individual Sports					.28
No. of Service Clubs	.2132			.25	
No. of Interest Clubs	.1702	.3109		.26	(-.01)
Trait Anxiety Level				-.36	-.32
Gordon Personal Profile					
Sociability trait					-.27
Emotional stability				.30	
Ascendency trait	.2218			.28	
Fatigue Symptoms-Trait	-.3193			-.41	
Fatigue Symptoms-State				-.29	-.26
Occupational Opinion					
AF policy/practices					.43
Assignment locality	.1933			.35	.43
Social status					.36
Organizational Climate					
Structure		.3272			(.09)
Risk	-.2241			(-.11)	
Warmth				.27	
Conflicts		-.3353			(.07)
Group Morale					
Satis. indiv. motives				.34	.29
Homogeneity of att.	.2456			.30	.27
Satis. interpersonal rel.					.34
Satis. with leader					.33
LBDQ					
Persuasiveness					.28
Consideration		-.4230			(.15)
Motivation					
Job curiosity trait	HP(.01)		HP(.01)	.37	.53
Persistence trait	HP(.01)	.6173	HP(.01)	.31	.61
Prof. identification	HP(.01)		HP(.01)	.29	.42
Team attitude			HP(.01)		.41
Org. identification			HP(.01)	.32	.38
Self-starter trait					.43

TABLE 6

Statistical Results on Factors Affecting Maintenance Performance
Accuracy Based on Data Collected Across 140 SRAM Technicians
(Compiled from Sauer, Campbell, Potter, and Askren, 1977)

Predictor	Significant High Performer Variables and Multiple Regr. Coefficients		Correlation Coefficients Correl. R's ≥ .25	
	462X0(B14.3824)	463X0(B13.5316)	462X0	463X0
Years of Service				.34
No. of Re-enlistments				.26
Months in Career Field				.32
No. of Interest Clubs		.1991	.32	
No. of Indiv. Sports		.1588		(.18)
Gordon Personal Profile:				
Responsibility trait			.29	
Fatigue Symptoms-Trait	** HP(.01)		-.40	
Fatigue Symptoms-State			-.26	
LBDQ:				
Representation				-.36
Tolerance of freedom			.26	
Occupational Opinion				
AF policy/practices	** HP(.01)		.27	
Assignment locality	** .3540-HP(.01)		.39	
Pay and benefits	-.2338		(.18)	
Promotion opportunity	** HP(.01)		.28	
Organization Climate				
Responsibility				-.29
Rewards	-.3263		(.06)	
Warmth			.29	
Group Morale				
Satis indiv. motives	* .3817HP(.001)		.41	
Homogeneity of attitude			.32	
Satis. with leader		-.3111		(.00)
Motivation				
Job curiosity trait	** .8249-HP(.01)	** HP(.01)	.46	.57
Persistence trait	** HP(.01)	** .7897 HP(.01)	.37	.71
Prof. identification	** HP(.01)	** HP(.01)	.32	.51
Team attitude	** HP(.01)	** HP(.01)		.50
Organ. identification	** HP(.01)	** HP(.01)	.37	.45
Self-starter trait	** -.4368 HP(.01)	** HP(.01)		.54

* Significant at ($p < .001$) for high performers

** Significant at ($p < .01$) for high performers

	Coefficient	Mean Value H. P.	Weight
1. Job curiosity trait	.8249	72.4	59.72
2. Self starter trait	.4868	62.9	27.47
3. Assignment locality	.3540	15.5	5.48
4. Organization climate/rewards	.3263	15.0	3.92
5. Group morale satisfying to individual motives	.3817	9.1	3.47
6. Pay and benefits satisfaction	.2338	9.2	2.15

The above analyses used the mean values of the scores for high performers, extracted from the data runs for the Sauer, Campbell, Potter, and Askren study. This source material is being filed for future reference.

It is of interest to see whether the prioritization holds for the low performers. For the regression equation for performance speed (462X0 technicians), the priority list for low performers computes as follows:

	Coefficient	Mean Value LP	Weight
1. Ascendency trait	.2218	15.7	3.4823
2. Organizational climate/risk	.2241	13.0	2.9133
3. Group homogeneity of attitude	.2456	11.3	2.7750
4. Assignment locality	.1933	11.9	2.3003
5. Fatigue trait	.3193	5.0	1.5965
6. Number of extracurricular interest clubs	.1702	0.1	0.0170
7. Number of service clubs	.2132	0.0	0.0

For accuracy of performance for the low performers, the weighted ordering becomes:

		Coefficient	Mean Value LP	Weight
1.	Job curiosity trait	.8249	48.2	39.7602
2.	Self-starter trait	.4368	44.4	19.3939
3.	Assignment locality	.3540	11.9	4.2126
4.	*Organizational climate/reward*	.3263	8.5	2.7736
5.	Group morale satisfying to individual motives	.3817	5.7	2.1757
6.	Pay and benefits satisfaction	.2338	7.0	1.6366

From the above analyses, it may be observed that the ordering of predictive variables for accuracy of performance is identical for AFSC 462XO high performers and low performers. For speed of performance, however, group morale in the form of homogeneity of attitude moves from 2nd to 4th place for the low performers, who in turn weight organization climate-risk (safety hazards) as more important than for the high performers. The reader is reminded that the performance measures used in the above analysis were speed and accuracy of task performance by experimental subjects as rated by the supervisors. The human resources data came from questionnaires administered to the experimental subjects. The only human resources factor which was a significant predictor of both speed and accuracy ratings of performance on MSET Task 1101A (462XO technicians) was Assignment Locality.

For the 463XO technicians (Nuclear Weapons Specialists), the regression equations for speed and accuracy of performance produced the following significant variables:

Performance Speed	Performance Accuracy
Years of Service	No. of Individual Sports
No. of Interest Clubs	No. of Interest Clubs
Organization Climate/Structure	Satisfaction with Supervisor
Organization Climate/Conflict	Persistence Trait
LBDQ - Consideration	
Persistence Trait	

In the same study (Sauer, Campbell, Potter, and Askren, 1977) human performance variables were also correlated with actual task performance times, as reported out on the AFTO 349 forms. Such correlations were developed across 6 MSET tasks. Stepwise multiple regression analyses were then run with the results summarized in Table 7.

However, the MSET 1324 task is not shown since the correlation R's were low and the three factors selected were different from all other tasks. From Table 7, the following summary can be drawn based upon the number of tasks showing significant Beta weights (surfaced in multiple stepwise regression):

Factor	No. of MSET Tasks with Beta Weights
Months in Career Field (or SRAM)	3
Fatigue: Trait or State	3
Extracurricular Service Clubs	3
Organization Climate/Structure	2
Organization Climate/Standards	2
**Satisfaction with Interpersonal Relationships	2
Organization Climate/Identity	2
Pay and Benefits Satisfaction	1
Social Status	1
Organization Climate/Warmth	1

TABLE 7

Beta Weight Coefficients for Regression Equations Relating Human Resources Data to Actual Performance Time (Compiled from Sauer, Campbell, Potter, and Askren, 1977)

X Var.	Human Resource Factors	MSET Task Number: 1101 Load Launcher	1102 Unload Launcher	1322 Payload Mate	1323 Payload Demate	1325 Demate Stores
1	Months in SRAM		.3389		-.3791	
2	Months in Career Field					.4772
3	Occupational Opinion					
4	Pay and Benefits Satisfaction	.2708				
5	Social Status	-.2408				
	Organizational Climate					
6	Reward	.1649				
7	Warmth		-.6592			
8	Structure		.4911		.3887	
9	Standards					.4813
10	Responsibility				-.3111	-.2655
	Group Morale					
11	Homogeneity of Attitude		-.3861			
12	Satisfied/Interpersonal Relations	-.1405	.5145			
13	Satisfied/Individual Motives			-.3230		
	Motivation					
14	Organizational Identification		-.1554			-.5747
15	Professional Identification			-.4637		
16	Job Curiosity Trait				.3206	
	LBDQ					
17	Production Emphasis			.7409		
18	Structure				.3846	
19	Fatigue Trait or State			-.2203	.2404	.3721
20	Extracurricular Service Clubs		.4090	-.2681		-.3308
21	Interest Club Officer	.2516				

MSET Task	Predictive Equation for Performance Time
1101	$y = 61.2095 + .2708X_4 + .2408X_5 + .1649X_6 - .1405X_{12} + .2516X_{21}$
1102	$y = A_2 + .3389X_1 - .6592X_7 + .4911X_8 - .3861X_{11} + .5145X_{12} - .1554X_{14} + .4050X_{20}$
1322	$y = A_3 - .3230X_{13} - .4637X_{15} + .7409X_{17} - .2203X_{19} - .2681X_{20}$
1323	$y = A_4 - .3791X_1 + .3887X_8 - .3111X_{10} + .3206X_{16} + .3846X_{18} + .2404X_{19}$
1325	$y = A_5 + .4772X_2 + .4813X_9 - .2655X_{10} - .5747X_{14} + .3721X_{19} - .3308X_{20}$

y = time in minutes

*Organizational Identification (Mission)	1
Homogeneity of Group Attitudes	1
Group Morale Satisfying Individual Motives	1
*Professional Identification	1
*Job Curiosity Trait	1
Production Emphasis	1
Interest Club Officer	1
Responsibility Trait	1

* Highest weighted factor on an MSET task.

The Beta weights for a particular task can also be multiplied by the mean values of the predictor variables and thus prioritize the variables. For the MSET 1101A task (Load Launcher), the results are as follows using the mean values from the computer output data (performance measure-actual performance time on MSET task):

	Factor	Coeff.	Mean Value	Weight
1.	Group morale satisfying to interpersonal relationships	.1405	18.0920	2.5419
2.	Pay and benefits satisfaction	.2704	9.3068	2.5166
3.	Organization climate/reward	.1649	12.0245	1.9828
4.	Social status	.2408	6.6319	1.5970
5.	Interest club officer	.2516	.0245	.0016

Similar results for tasks 1102, 1322, 1323, and 1325 are shown in Table 8. Although the predictor variables for the five tasks vary somewhat in nature and in order, there are common trends apparent.

Errors cited in airmen following of technical orders were also analyzed in the study (10, 1977), as well as safety and reliability errors. Correlation analyses and regression analyses were run for data collected on nine MSET tasks. Since the study was primarily concerned with Nuclear

TABLE 8

Prioritized Predictive Factors for Time to Perform Five
Selected MSET Tasks Based on Regression Equations
(Data from Computer Runs Analyses by Sauer, Campbell, Potter, and Askren, 1977)

Priority	MSET Task 1101	MSET Task 1107	MSET Task 1322	MSET Task 1323	MSET Task 1325
1	Interp. Relations	Interp. Relations	Prof. Identification	Job Curiosity Trait	Org. Identification
2	Pay & Benefits Satis.	Org. Identification	Production Emphasis	Leadership/Structure	Mos. in Career Field
3	Org. Climate/Reward	Warmth of Org Climate	Satis Indiv Motives	Months in SRAM	Org./Standards
4	Social Status	Org./Structure	Fatigue Trait	Org./Structure	Org./Responsibility
5	Int. Club Officer	Homo. of Gr. Attitude	Service Clubs	Org./Standards	Fatigue Trait
6	---	Months in SRAM	---	Fatigue State	Service Clubs
7	---	Service Clubs	---	---	---

Weighting Calculations-Tasks 1102-1325

Task	Factor	Coeff.	Mean Value	Weight	Task	Factor	Coeff.	Mean Value	Weight
1102	Interp. Relations	.5145	18.322	9.4267	1323	Job Curiosity Trait	.3206	63.545	20.3725
	Org./Identification	.1554	56.271	8.7445		Leadership/Structure	.3846	37.836	14.5517
	Org./Warmth	.6592	11.864	7.8207		Months in SRAM	.3791	19.309	7.3200
	Org./Structure	.4911	11.593	5.6933		Org./Structure	.3887	16.108	6.2612
	Homogeneity	.3861	12.305	4.7510		Org./Standards	.3111	15.182	4.7231
	Months in SRAM	.3389	13.542	4.5894		Fatigue State	.2404	2.727	.6556
	Service Clubs	.2681	.396	.1062					
1322	Prof. Identification	.4637	55.472	25.7224	1325	Org./Identification	.5747	46.182	26.5408
	Prod. Emphasis	.7409	16.528	12.2456		Mos. in Career Field	.4772	24.364	11.6265
	Satis. Ind. Motives	.3230	5.906	1.9076		Org./Standards	.4813	13.164	6.3358
	Fatigue Trait	.2203	2.717	.5986		Org./Responsibility	.2655	15.182	4.0308
	Service Clubs	.2681	.396	.1062		Fatigue Trait	.3721	2.600	.9675
						Service Clubs	.3308	.400	.1323

missile handling safety, and the correlation and multiple regression correlation coefficients for safety errors were low and mostly non-significant, the analysis on safety and reliability errors was not included in the report. However, the computer runs have been analyzed by this writer for errors reported in following technical orders across the nine MSET tasks. Table 9 lists the human resource factors which surfaced in the regression studies and the number of tasks in which each factor surfaced. The result is another ordering of human resource factors affecting performance accuracy.

One task of the nine had the largest number of entries (195) on technical order errors. Prioritization of significant factors for this task (MSET Task 4) is shown below:

Factor	Beta Coefficient	Mean Value	Weight
Age (AFSC level)	.2229	21.5254	4.7980
Social Status	.2904	6.6949	1.9448
*Success Level in Academic Major	.4691	1.8983	.8905
Number of Interest Clubs	.2516	.2034	.0512
Number of Service Clubs	.3326	.0678	.0226
	Multiple Regression R =		.6764

* Eliminated from latter analyses

For another of the tasks, only Months in Career Field surfaced as the predictor with a multiple regression R = .2570

Summary on Significant Equipment/Information Factors and Human Resource Factors Affecting Maintenance Performance

From the body of research performed to date, it is possible to hy-

TABLE 9

Prioritized Predictive Factors for Performance

Accuracy as a Function of Technical Order Errors

Across Nine MSET Tasks on SRAM

(Data from Computer Runs Analyses by Sauer, Campbell,

Potter, and Askren, 1977)

Predictive Factor	No. of Tasks Factor Surfaced in Regression Equations
Months in Career Field	3
Months in SRAM	3
Pay and Benefits Satisfaction	3
Age or AFSC Level	3
Number of Service Clubs	2
Number of Interest Clubs	2
Service Club Officer	2
Years of Service	1
Number of Re-enlistments	1
Rank	1
Fatigue Trait	1
LBDQ: Initiation of Structure	1
Representation	1
Assignment Locality	1
Social Status	1
Occupational Opinion Overall	1
Occupation Climate/Responsibility	1
Group Morale: Satisfaction with Leader	1
Homogeneity of Attitude	1
AQE Score	1
Motivation: Persistence Trait	1
Professional Identification	1

Total Population for Data: 629 Entries

pothesize on those factors which most influence the speed and accuracy of maintenance performance in the USAF. These factors have proven important for selected types of maintenance activity, and it seems likely that the same factors may influence maintenance performance effectiveness in general across the Air Force. Testing of this hypothesis is to be proposed. Although some of the following factors appear to most affect performance speed and others performance accuracy, there is considerable commonality of performance influence; thus maintenance performance effectiveness is considered in toto as a combination of speed and accuracy.

Equipment/Information Factors

The most important equipment/information factors in the Taxonomy appear to be:

Maintainability

- Weight/size of subsystem and/or component
- Access for test and check
- Clearance for removal/replacement
- Number, type, and arrangement of internal components

Reliability

- Design operating parameters of equipment
- Average operating hours between services

Testing and Repair

- Type and amount of test-equipment information provided
- Adequacy and completeness of technical information on equipment
- Level of information received from tests (and test equipment)

Human Resource Factors

The most important human resource factors in the Taxonomy appear to be:

Experience

AFSC level (or Age)

Months in career field

Years of service

Skills

For test equipment used

Knowledge

Of maintenance procedures

Organizational Climate and Group Morale

Homogeneity of group attitude

Organization climate/warmth

Organization climate/risk

Organization climate/structure

Satisfaction of individual motives

Organization climate/responsibility

Satisfaction of interpersonal relations

Satisfaction with leader

Personal Traits and Motivation

Organizational identification

Fatigue trait

Self-starter trait

Job curiosity trait

Ascendency trait

Persistence trait

Occupational and Environmental Conditioners

Participation in interest or service clubs

Participation as an interest or service club officer

Pay and benefits satisfaction

Assignment locality

Social status satisfaction

The single most important "equipment/information" predictor of performance rate appears to be the adequacy and completeness of the technical information provided on the equipment (TO's, JGM's, etc.). The most important "human resources" predictors of performance rate appear to be: (1) airman motivation, (2) organizational climate and group morale, (3) months in career field, and (4) assignment locality.

For accuracy of performance, the human resource predictors which appear to be most important are: (1) airman motivation, (2) assignment locality, (3) months in career field, and (4) pay and benefits satisfaction.

A factor which has not been evaluated is the impact of required non-maintenance activities which detract from maintenance time availability.

A GENERALIZED MODEL OF MAINTENANCE MANPOWER EFFECTIVENESS

A generalized model of base maintenance manpower effectiveness for a given maintenance squadron, a given end item equipment, and a given time period can now be postulated as:

$$y = a_1x_1 + a_2x_2 + a_3x_3 + \dots a_ix_i \dots + a_{35}x_{35} + a_{36}x_{36}$$

where y = Performance index for maintenance manpower effectiveness

a_i = Coefficients or multipliers, $i = 1, 2, 3, \ldots 36$

and x_1 = Average weight of subsystems/components serviced by squadron

x_2 = Average accessibility/clearance rating for test, check, remove, and/or replace of subsystems serviced

x_3 = Average number of internal components in subsystems serviced

x_4 = Average operating hours between maintenance actions for subsystems serviced

x_5 = Average supervisor rating of the adequacy and completeness of technical information on the subsystems serviced and on the test equipment used

x_6 = Average supervisor rating on the test equipment usability and value of test output for the subsystems serviced

x_7 = Average supervisor rating of the knowledge and skills of technicians within the squadron for the subsystems serviced

x_8 = Mean outside temperature during evaluation period

x_9 = Average number of years of AF service in squadron

x_{10} = Average number of months of service in assigned AFSC's, within squadron

x_{11} = Average primary AFSC levels in squadron

x_{12} = Average number of extracurricular service and interest clubs participated in by technicians in squadron

x_{13} = Average technician satisfaction with supervisor on Goldman Group Morale Scale

x_{14} = Average technician satisfaction with group objectives relative to individual motives on Goldman Group Morale Scale

x_{15} = Average technician feelings of homogeneity of group attitude on Goldman Group Morale Scale

x_{16} = Average technician satisfaction with interpersonal relations on Goldman Group Morale Scale

x_{17} = Average technician feelings of organization risk on Organization Climate Inventory

x_{18} = Average technician feelings of organization reward on Organization Climate Inventory

x_{19}= Average technician feelings of organization structure on Organization Climate Inventory

x_{20}= Average technician feelings of organization conflict on Organization Climate Inventory

x_{21}= Average technician feelings of organization identity on Organization Climate Inventory

x_{22}= Average technician feelings of organization warmth on Organization Climate Inventory

x_{23}= Average technician satisfaction with assignment locality on the Occupational Attitude Inventory

x_{24}= Average technician satisfaction with social status of job on the Occupational Attitude Inventory

x_{25}= Average technician satisfaction with pay and benefits on the Occupational Attitude Inventory

x_{26}= Average technician fatigue trait on the modified Yoshitake Fatigue Symptoms Checklist

x_{27}= Average technician professional identification on the Motivation Index as rated by supervisors

x_{28}= Average technician organizational identification on the Motivation Index as rated by supervisors

x_{29}= Average technician job curiosity trait on the Motivation Index as rated by supervisors

x_{30}= Average technician self-starter trait on the Motivation Index as rated by supervisors

x_{31}= Average technician persistence trait on the Motivation Index as rated by supervisors

x_{32}= Average technician ascendency trait on the Gordon Personal Profile

x_{33}= Average technician responsibility trait on the Gordon Personal Profile

x_{34}= Average technician emotional stability trait on the Gordon Personal Profile

x_{35}= Average technician sociability trait on the Gordon Personal Profile

x_{36}= Average technician estimate of percent of time spent on non-maintenance activity

The values for the x_i's would come from the following sources:

x_1 -x_3	:	Equipment Design Specifications
x_4	:	Logistics data via PCN-D056B logs
x_5 -x_7	:	Supervisor ratings using scales to be designed
x_8	:	Climatological data
x_9 -x_{12}	:	Biographic information form to be designed-Technician
x_{13} -x_{16}	:	Goldman Group Morale Scale (Goldman, B., Group Cohesiveness: A Study of Group Morale, Manual of Instructions. Roosevelt University, 1958) - Technician
x_{17} -x_{22}	:	Organization Climate Inventory (Litman, G.H. and R.A. Stringer, Motivation and Organizational Climate, Harvard University, 1968) - Technician
x_{23} -x_{25}	:	Occupational Attitude Inventory (Tuttle, T.C., R.B. Gould and J.T. Hazel, Dimensions of Job Satisfaction: Initial Development of the Air Force Occupational Attitude Inventory, AFHRL-TR-75-1, ADA 014796, Lakeland AFB, June 1975) - Technician
x_{26}	:	Modified Yoshitake Fatigue Symptoms Checklist, (Yoshitake, H., Relations Between the Symptoms and the Feeling of Fatigue." In K. Hashimoto, K. Kogi, and E. Grandjean (eds.), Methodology in Human Fatigue Assessment, Taylor and Francis, London, 1971). - Technician
x_{27} -x_{31}	:	Motivation Index (Landy, F.J. and R.M. Guioni, "Development of Scales for the Measurement of Work Motivation," Organizational Behavior and Human Performance, 1950, 5, 93-103) - Supervisor
x_{32} -x_{35}	:	Gordon Personal Profile (Gordon, L.V., Gordon Personal Profile: Manual, Harcourt, Brace, and World, 1963) - Technician
x_{36}	:	Information item to be designed - Technician

The model should be applicable to both base level maintenance and depot maintenance. However, certain terms could drop out (have zero coefficients) for a particular level of maintenance. In the case of engine depot maintenance, it was previously shown that engine dimensions and

engine operating parameters also impact upon maintenance performance. The purpose of the model is not to predict the amount of maintenance required, but rather to either predict or evaluate the effectiveness of technician maintenance performance when servicing is required.

The a_i coefficients need to be determined by further research. It can be expected that the set of coefficients will be different for different end-item equipments, different Commands, and different maintenance squadrons. What is needed for use of the predictive/evaluative model are the set of coefficients which are applicable and the average values of the 36 performance drivers for a particular type of maintenance squadron at an evaluative point of time.

EXPERIMENTAL DESIGN FOR FURTHER RESEARCH

Past research studies have identified the significant predictive drivers of performance, as measured by speed and accuracy of work (or time to perform and number of technical errors identified). Such past studies have also generated sets of regression equation coefficients for particular equipment subsystems, AFSC's, and tasks, as presented in this paper. Sample sizes were relatively small.

A large-scale research study is proposed to accomplish the following purposes:

1. Further test the importance or significance of the proposed performance drivers across a larger sample of AF equipments, AFSC's, and maintenance levels/tasks.

2. Generate the a_i coefficients applicable to particular end-item equipments (and subsystems) across the four types of base maintenance squadrons: organizational, field, avionics, munitions.

3. Establish normative values for the x_i drivers and for the overall performance measure, y, for each end-item equipment and type of maintenance squadron.

4. Prioritize further the key elements (drivers) affecting maintenance performance effectiveness.

5. Generate a table of formula coefficients applicable to particular maintenance squadrons and end-item equipments, and prepare an operational manual for use of the model. The manual would include normative values of the driver variables x_i and performance measure y.

In the long run, complete validation of the model and generation of coefficients will necessitate collection of data across the Air Force. However, it would seem that the place to start is with one Command, such as SAC. If SAC were selected, the following end-item equipments could be included: B52, FB111, KC135, RF4. These equipments would incorporate most of the types of subsystems requiring AF servicing.

The data sample should encompass within the Command all end-item equipments listed, all appropriate subsystems, all AFSC maintenance specialty codes which are applicable, and the major action taken categories. At least three Command bases are suggested as data collection sources.

The performance data sources would be as follows:

1. Speed of technician performance: supervisor rating of each technician in the squadron on overall performance capability relative to rate of work.

2. Accuracy of technician performance: supervisor rating of each technician in the squadron on overall performance capability relative to accuracy (quality) of work.

3. Average performance time: LOG-MMO (AR) 7169 of AFLC summarizes, on demand and by Work Unit Codes, total number of man-hours expended during the most recent 12-month period and the number of maintenance actions (units worked on) for same period, per Command and Command Base. Gives average hours expended per repair on every appropriate work unit code for each end-item equipment serviced at each Base. The data is separated into on-equipment actions and off-equipment (or shop) actions, but does not include depot maintenance. However, the data is not separated by maintenance squadron. Therefore, some effort will be needed to ascertain for particular WUC's, HOWMALS, and Action Taken Codes what squadron normally would perform the maintenance activity. Data is based on AFTO 349 form.

4. Average number of performance errors: A statistical review of Command Summaries of the MSEP evaluations will provide this information over a most recent 12-month period. Data is based on AF 2416 form and others.

One intent of the research would be to compare measures 1 and 3 above and measures 2 and 4 above, to evaluate the level of agreement

between averaged supervisory ratings of performance speed and accuracy and the reported data submitted on AFTO 349 and AF 2416. If there is good agreement, then either the reported data or the supervisory ratings can be used in future research. If the level of agreement is low, then additional research is needed on whether the subjective or objective data best represents real world experience.

The performance variable proposed for this research is a combination of the average speed and accuracy of squadron technician performance. It is suggested that the quality of performance may be more critical than the speed of performance and should be weighted more heavily in the combined criterion variable. For instance, if both averaged speed and averaged accuracy performance ratings are converted to percent, the measure of combined performance variable y might be:

$$y = w_1 + 1.2w_2$$

where y = combined percentile effectiveness rating

w_1 = percentile rating of averaged speed of performance by technicians in squadron across all Action Taken Codes.

w_2 = percentile rating of averaged accuracy of performance by technicians in squadron across all Action Taken Codes.

If <u>time to perform</u> and <u>number of errors</u> are used in the combined measure, based on the objective or reported data, a different conversion to percentiles would be needed but the percentile addition equation would be the same.

The 36 predictor variable measures necessitate data collection via supervisor and technician questionnaires, climatological data from the U.S. Weather Bureau or Geographical Survey, logistics data via Logs produced by the AFLC, and equipment design specifications via AFFDL. It should be noted that not all of the questions in the standardized questionnaires are needed to provide input values for the 36 predictor variables; where this is so it is suggested that the pertinent items be lifted and incorporated into an overall technician questionnaire comprised as follows:

1. Biographic information, including number of years of AF service, number of months of service within current AFSC, primary AFSC level, number of interest and/or service clubs participating in, and technician's estimate of percent of time spent on non-maintenance activities.

2. Organization Climate Inventory items covering risk, reward, structure, conflict, identity, and warmth.

3. Occupational Opinion Survey items covering assignment locality, social status, and pay and benefits.

4. Modified Yoshitake Fatigue Symptoms Checklist.

5. Gordon Personal Profile items covering ascendency trait, responsibility trait, emotional stability trait, and sociability trait.

6. Goldman Group Morale Scale items covering satisfaction with supervisor, satisfaction with group objectives meeting individual

motives, homogeneity of group attitude, and satisfaction with interpersonal relationships.

The proposed study would involve 6 man-months of senior researcher time, 3 man-months of junior researcher time, support for computer analyses and for LOG data reports from AFLC, and support from one or more AF Commands. The Command Support would consist of providing technician and supervisor time allowances for project orientation and for completion of questionnaires and evaluations. Approximately one hour would be required from each technician, and a sample size of 1000 technicians may be needed.

The study procedure would include:

1. Generation of survey instrument for technician completion incorporating needed biographic data and extracted items from the several evaluation scales.

2. Generation of technician evaluation instrument for supervisor completion.

3. Selection of Command(s) and Bases for data collection.

4. Collection of 12-months Climotological temperature data for Bases selected.

5. Collection of needed design and operating specifications for end-item equipments appropriate to Command and Bases.

6. Collection of 12-months of logistics LOG data on end-item equipments and WUC's.

7. Collection of 12-months history of MSEP reports from Command headquarters, by end item equipment and Base of operation.

8. Experimental design activity for sampling process and data handling. Trials of the survey instruments.

9. Schedule of orientation meetings at Bases and for administration of questionnaires.

10. Conduct of orientation meetings at Bases and administration of questionnaires.

11. Programming of computer analysis techniques (or collection of existing packages into a unified data evaluation system)

12. Composing data and keypunching operations

13. Computer statistical runs

14. Evaluation of statistical reports

15. Drawing of conclusions and preparation of project report

16. Final briefing of AFHRL

REFERENCES

Askren, W. B., W. B. Campbell, D. J. Seifert, T. J. Hall, R. C. Johnson, and R. H. Sulzen, Feasibility of a Computer Simulation Method for Evaluating Human Effects on Nuclear Systems Safety, AFHRL-TR-76-18, AFWL-TR-76-15, Air Force Human Resources Laboratory, WPAFB, 1976.

Beu, N. J. and R. C. Nichols, More Maintenance in OMS, Air University Report, Maxwell AFB, ADB 019707, 1977.

Carpenter-Huffman, P. and B. Rostker, The Relevance of Training for the Maintenance of Advanced Avionics, Rand Corporation Report R-1894-AF, AD 047707, 1976.

Drake, K. L., M. S. Sanders, W. H. Crooks, and G. Weltman, Comparative Studies of Organizational Factors in Military Maintenance, Perceptronics Corporation Report PTR-1043-77-10, ADA 050403, 1977.

Foley, John P. Jr., Some Key Problems Concerning the Specification, Development, and Use of Task Identification and Analyses, AFHRL-TR-76-57, Air Force Human Resources Laboratory, WPAFB, 1976.

Guthrie, W. J., A Study of Job Characteristics and Job Attitudes at a Tactical Air Command Fighter Complex, Master's Thesis, School of Engineering, AFIT/GSM/SM/77S-5, ADA 045998, 1977.

LaRue, R. D. and G. D. Metzinger, Motivation of Maintenance Personnel to Work Particular Shifts, Master's Thesis, School of Logistics Management AFIT/SLSR 15-77A, ADB 019707, 1977.

Meister, D., D. L. Finley, and E. A. Thompson, Relationship Between System Design, Technician Training, and Maintenance Job Performance on Two Autopilot Subsystems, Bunker-Ramo Corporation Report AFHRL-TR-70-20, Air Force Human Resources Laboratory, WPAFB, 1971.

Pritchard, R. D. and R. V. Montagno, Effects of Specific vs Non-specific and Absolute vs Comparative Feedback on Performance and Satisfaction, AFHRL-TR-78-12, Air Force Human Resources Laboratory, WPAFB, 1978.

Wiley, L. L. and C. P. Hahn, Task Level Performance Criteria Development, AFHRL-TR-77-75, Air Force Human Resources Laboratory, WPAFB, 1977.

APPENDIX

Composite Missile Technician Rankings of Factors Affecting Maintenance Performance

TABLE 10

Composite Rankings of the Contributions of Airman Resource Factors and Equipment/Environment Factors to Maintenance Performance Effectiveness (Compiled from Sauer, Campbell, Potter, and Askren, 1977)

Key Human Resources Factors[1]

	Missile System			Career Fields			Bases					
Ranked Human Resource Factors	SCRAM	GENIE	Minuteman	462XO	463XO	443XO	1	2	3	4	5	Composite Rank
Team Cohesiveness	1	1	3	1	1	1	1	1	1	1	1	1
Fatigue	3	5	1	3	2	3	4	3	3	3	3	2
Systems Training and Experience	4	3	2	5	3	2	2	2	4	4	7	3
Emotional Stability	2	2	6	2	4	6	6	7	2	2	2	4
Leadership	5	8	5	4	5	7	3	4	6	5	4	5
Motivation	6	4	4	6	6	4	5	5	5	7	8	6
Career Field Training	7	6	7	7	7	5	7	6	7	6	6	7
Military Morale and Attitude	9	10	8	8	10	9	9	8	8	10	9	8
Organizational Structure	8	9	9	9	8	10	8	9	10	8	5	9
Aptitude	10	7	6	10	9	8	10	10	9	9	10	10

[1] Factor ranked 1 had most influence

TABLE 10 (Continued)

Key Equipment/Environmental Factors[1]

	Missile System			Career Fields			Bases					
Ranked Human Resource Factors	SCRAM	GENIE	Minuteman	462XO	463XO	443XO	1	2	3	4	5	Composite Rank
Equipment Reliability	1	1	1	2	1	2	1	1	1	1	3	1
Weather Conditions	2	3	2	1	4	1	2	2	2	2	2	2
Operation of Equipment	3	4	4	3	3	4	4	3	4	5	1	3
Technical Orders	5	7	3	7	2	3	3	4	6	7	4	4
Lighting Conditions	6	2	5	4	7	5	5	5	7	3	7	5
Noise Level	4	5	8	5	6	9	7	7	3	4	6	6
Equipment Safety Features	7	8	7	6	8	6	8	6	5	6	5	7
Workplace Size and Shape	8	6	6	9	5	7	6	8	8	8	8	8
Clothing Types	9	9	9	8	9	8	9	9	9	9	9	9

AIR FORCE LOGISTICS COMMAND

Research Associates:

Michael J. Cleary, Wright State University

Gordon K. Constable, Wright State University

AN INVESTIGATION OF PRODUCTION LEADTIME FORECASTING FOR AIR LOGISTICS CENTERS

BY

MICHAEL J. CLEARY

12 JUNE 1978 - 18 AUGUST 1978

Prepared For

USAF-ASEE Summer Research Faculty Program
Air Force Office of Scientific Research
Washington, D. C.

And

Directorate of Management Sciences (XRS)
Deputy Chief of Staff/Plans and Programs
Headquarters Air Force Logistics Command
Wright-Patterson Air Force Base, Ohio 45433

AN INVESTIGATION OF PRODUCTION LEADTIME

FORECASTING FOR AIR LOGISTICS CENTERS

Michael J. Cleary

Wright State University, Dayton, Ohio

Abstract

Approximately four years ago Procurement became concerned with variations in Production Leadtime (PLT). The origin of this concern was the great increase in overall PLT from early 1973 through 1975. These increases were caused by a series of economic factors such as price controls and inflation, which in turn resulted in a general unstable economic setting. Within this economic climate great demands arose in various sectors of industry, and in turn lead to increases in PLT. This report investigated prior research on PLT and developed a specific forecasting model which attempted to predict PLT for the materials handling industry. Conclusions indicated that a macro external forcasting model would be ineffective in an operation setting due to problems in dissaggregation down to individual items. Further, it was concluded that some improvements could be made over the present system of using the last PLT as an estimate of the future PLT. One possibility would be to investigate the usefulness of integrating a smoothing technique for individual items into the inventory data system.

PREFACE

This study was undertaken as part of the USAF-ASEE Program during the period 12 Jun 1978 to 13 Aug 1978. The report investigates the possibility of developing a forecasting model for production leadtime.

The author of this report extends thanks to Dr. Dickison and others of XRS who gave various support in the preporation of this report.

TABLE OF CONTENTS

CHAPTER I

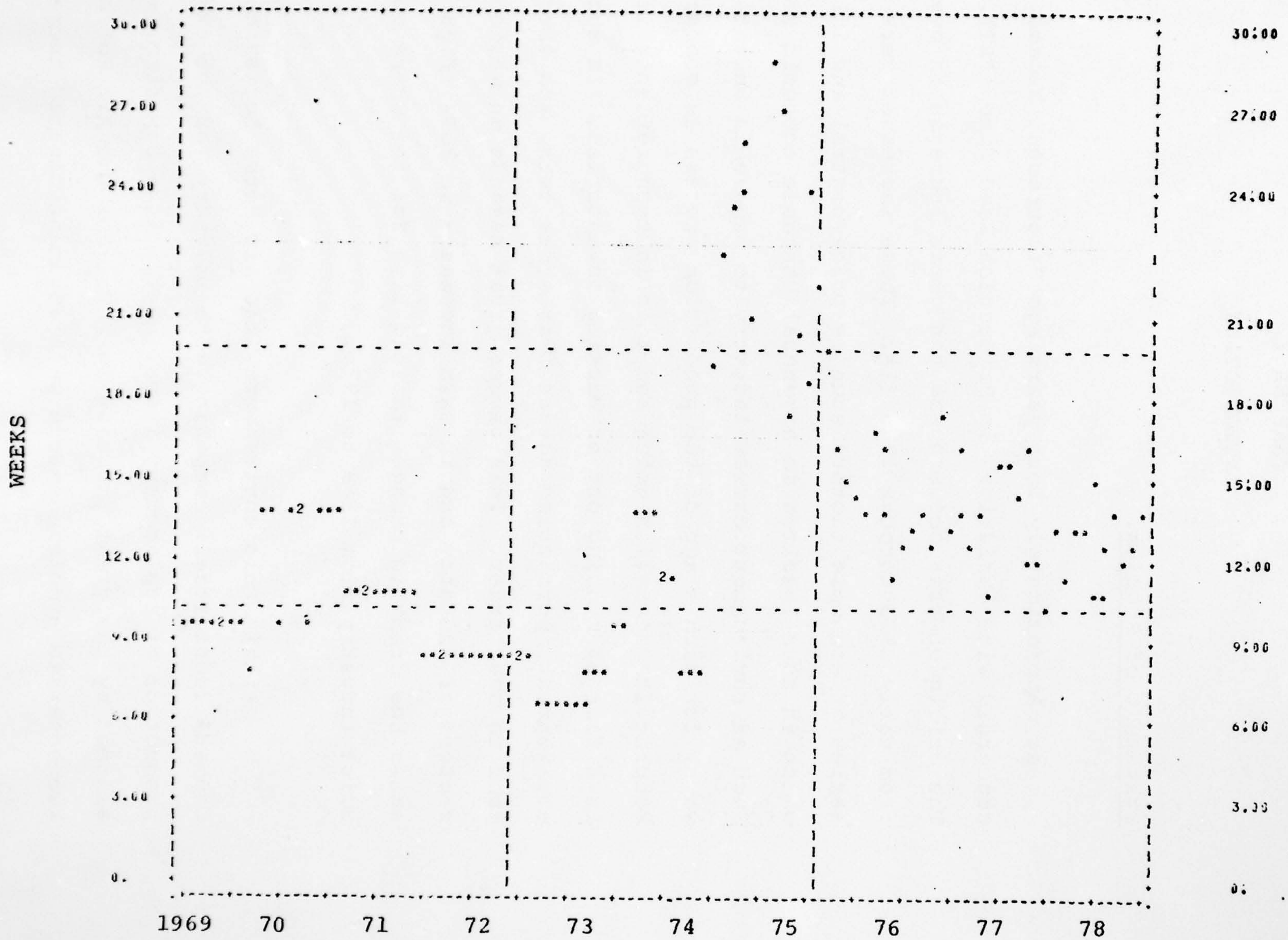
INTRODUCTION

Statement of Problem.

a. Approximately four years ago Procurement became concerned with variations in Production Lead Time (PLT). The origin of this concern was the great increase in overall PLT from early 1973 through 1975 (1). These increases were caused by a series of economic factors such as price control and inflation, which in turn resulted in a general unstable economic setting. Such an environment causes industry to behave in unwanted ways, for each stage of the production process is concerned with keeping its operation going and this in turn reflects itself in a fear of running out of working inventories. In such an environment, purchasing agents fear stock outs, and therefore tend to over order. This causes great demands on various sectors of industry and in turn increases in PLT. Figure 1.1 shows the dramatic changes in lead time for the Crane and Hoist Industry from 1969 to 1978.

b. Within this environment, the Air Force Logistics Command found itself competing with industry for the expeditious production of replacement items. They, like any firm, were alarmed by the great increase in PLT. For the Air Force, two separate problems emerge with an unanticipated increase

FIGURE 1.1
CRANES & HOISTS



AD-A065 651 OHIO STATE UNIV RESEARCH FOUNDATION COLUMBUS F/G 1/3
USAF-ASEE (1978) SUMMER FACULTY RESEARCH PROGRAM (WPAFB). VOLUM--ETC(U)
NOV 78 C D BAILEY F44620-76-C-0052

UNCLASSIFIED AFOSR-TR-79-0232 NL

4 OF 5
AD
A065651

in PLT. The first of these is that as the pipeline of parts dries up, inventory stock-outs occur. This causes the level of readiness of the Air Force to decrease. The second problem is budgeting. In order to deal with replacement parts for the Air Force, Procurement is given a budget to finance its stock fund. Each year, estimates of the number and type of buys for all Air Force products are made. An important input into that process is a separate estimation of PLT. If that estimate is inaccurate, total buy replacement dollars are insufficient. This is a result of the increased cost when PLT increases, because of the associated inflation. DCS/Procurement and Production has therefore developed the following plan to deal with this problem.

Procurement and Production Plan.

Production and Procurement decided to attack this problem at two different levels: First, by developing a method which would update PLT on individual items and second, by utilizing a model which could predict changes in PLT for industrial groupings. It was also hoped that at some time the two models could be interfaced, possibly integrating them through the computer system D041 and D062, which initiates the buy orders.

These orders are partially driven by PLT. The first step taken by Procurement was the initiation of a PLT update system for all Air Logistics Centers (ALCs). It was hoped that this system would improve the present rule for updating PLT which is to use the actual PLT from the last buy as the forecast for PLT on the next buy. (AFLC Regulation R 57-6, 29 September 1977 page 1-3).

> "Production Lead Time (PLT). Periods in time in whole months between date of contract or purchase order award and receipt of the first significant delivery quantity (under normal delivery conditions) based on the lastest contract or purchase order or contract."

This information is manually inputted into the DO42, (reparable items) and will soon be automatically inputted into the DO62 System (EOQ-or throw away items). Item managers can override this value if they have reason to believe that the PLT of the last buy is not a good estimate of the future PLT. The issue of using prior PLT as an estimate of future PLT is debatable and will be discussed later, but it seemed obvious to Procurement that the greatest errors in PLT would be on items that have not been purchased for a long period of time, since PLT can vary greatly, with various economic conditions.

Because of this, the first step taken was to develop a system which would identify a list of items that would be in a buy position in the next year, which also had not been purchased for the last six months (NOTE: See Appendix A for AFLCR 84-4 which defines the processs). This list is automatically generated from the D041, D062, J041 and J014 System which includes both reparable and EOQ items. The list is generated in March based on a December 31st computational process. A form letter is then sent out for each item to the last contractor, requesting an estimate of the present PLT for that item. Contractors have no obligation to respond to this letter, but the response has been about 70% for the last three years. These letters are then sent to item managers who are to change the PLT in the system (K/A file maintenance) if the PLT estimated by the contractor is substantially different from what is presently in the system.

The second step taken by Procurement in securing accurate estimates was the development of a predictive model for PLT which is independent of actual items purchased by the Air Force. The System started with the tracking average PLT as defined by <u>Production Magazine</u>, for various industrial groups. This is updated on a monthly basis and is used as a device to keep Procurement posted of any possible increases or decrease

in PLT. In addition to the tracking of industrial groups, Procurement began the development of a regression model which used a series of independent variables such as backlogs, new orders, etc., to predict PLT. At this point, XRS was requested to complete this model and the Summer Faculty project given to me was to continue the work of developing a predictive model for Production Leadtime.

CHAPTER II

SURVEY OF LITERATURE

A great amount of research on leadtime has been undertaken both in and out of government. One typically defines leadtime as being separated into the production and administrative aspects of the purchasing process. Administrative leadtime is defined as the period of time from the existence of a demand until a purchase order is sent or a contract signed. From that time, until the first significant delivery is received (at least 10%) , is defined as Production Leadtime. Since the military has organizational control over Administrative Leadtime, more research has been done in this area than in the area of Production Leadtime. In order to gain an overall view of what has been done in the area of Production Leadtime, one can first look at some basic discriminating factors which have lead to various types of leadtime research. This report will then examine a few specific models within these categories so as to gain an understanding of the present state-of-the-arts for the prediction of Production Leadtime.

Areas of Research:

a. Base Versus Air Logistics Command (ALC):

The most fundamental discriminating factor of PLT research is the level at which the procurement takes place.

The Air Force orders items at two levels, the Base and Air Logistics Command. These two levels face entirely different problems. At the ALC there are a number of item managers,
1200
each of whom has responsibility for 200 to 1,200 individual items. Those that have a fairly small number of items to manage, either have items that are ordered often, or have high dollar values. The item manager has a great deal of information about each item which is generated by the D041 or D062. Both systems include the current estimate for PLT which is unique for each item. As described in Chapter I, PLT can be changed on the basis of the contractor's survey, automatically changed to the last buy through the data system, or manually changed by the item manager based upon his personal judgment.

At the Base level, computation and implementation of PLT is entirely different. Each Base takes the previous year's data for PLT on all items purchased (primarily locally-purchased EOQ items) and uses this as the average PLT for all items. In addition, all items that had an actual PLT greater than 79 days are treated as outliers, and therefore not used in the computation of the average base PLT. This estimate of PLT is used for one year, at which time a new estimation is made based on the current year's data. The person administering the purchase is not an item manager, but rather a base supply manager. He has no way of overriding the calculated average PLT, even if his personal judgment indicates that the estimate is a poor one.

At most bases, an order is initiated automatically using some sort of classical EOQ base model, an input of which is the estimated PLT. This of course leads to extremely poor forecasting, since the items purchased have large deviations in Production Leadtime, yet the system assumes each product has exactly the same leadtime. The above description of two operations should suggest the fact that the problems, and therefore research associated with the two levels, are extremely different. At the base level, the main concern is identifying those items that traditionally have much longer PLTs than the base average for all items. Much of the research is focused on the shape of the distribution of PLT at the various bases (2). The second aspect of this research is identifying managerial methods of handling items that have traditionally had long PLTs. Current research in this area is pursuing a method which would have a series of PLTs that could be assigned to various item groups. It is believed that such a method would be more reflective of actual PLT than the present overall average method of prediction.

At the ALC level, one can assign unique PLTs for individual items. Therefore, the research has focused on methods that would allow one to externally establish the appropriate PLT and then assign that value to individual items (7).

b. Economic Ordering Quantity Versus Reparable Items:

Another difference in the procedure and research efforts is the type of item purchased. The Air Force distinguishes between reparable items and expendables (known as EOQ). The item managers at the ALC facility are categorized in terms of responsibility for either reparable or EOQ items. Further, AFLC has developed separate data systems, the D041-reparable and D062-EOQ, for the two categories of items.

The importance of PLT in initiating buy orders seems to be greater for EOQ than for reparable items. This is due to the fact that the most sensitive parameters associated with buy orders for reparable items seem to be the average turn-around time on repairing items and the estimated condemnation rate. By the same token, the reparable items typically have a much higher unit cost, and therefore mistakes in estimated PLT can have significant effects upon the total buy dollars for AFLC. Another dimension of the dichotomy between these two groups is that the manager of reparable items typically has a better knowledge of each item, since he has fewer items to manage. Therefore, there is a greater opportunity to have some sort of computerized estimating procedures for the EOQ-item manager, than for the reparable-item manager.

c. Tracking Versus Predicting:

The next dichotomy of type of research is the form of the model one would consider in estimating future PLT. One approach would be to look at models that simply smoothe out variations in PLT history and, in turn, use this smoothed value as a predictor for future production leadtime (4). A second method would be to develop an external forecast of future production leadtime based on a model which utilizes outside information such as economic data.

The summer faculty project viewed production leadtime from the vantage point of the ALC, for both reparable and EOQ items, utilizing both smoothing and external forecasting methods. The remainder of this chapter will deal with a review of various models that that dealt with production lead-time.

Prior PLT Prediction Models:

a. Westinghouse Model:

Like most large firms, Westinghouse was affected by the boom-bust period of 1973-75 when material shortages were followed by large inventory surpluses. The company developed a model which would predict PLT for the Hot and Cold Rolled Steel Industry. This commodity group was selected because the

general nature of that area seemed to provide a good barometer for the PLT of products they purchased. Overall findings were that manufacturers' inventories and capacity utilization were drivers of PLT for the Hot and Cold Rolled Steel Industry.

The actual model developed by Westinghouse was a regression model using ordinary least squares. The dependent variable was PLT for the Hot and Cold Rolled Steel Industry, the source of which was Purchasing Magazine. This journal surveys 7,000 of its readers twice a month for an estimate of current PLT. Persons responding to the survey are purchasing agents throughout the country. When Purchasing Magazine began the service, they noticed that various products have extremely different production leadtimes. Because of this, they have chosen 125 product types (see Table 2.1) on which to report. Secondly, they found that when they looked at any individual item, the responses of the purchasing agents for average leadtime did not follow a normal distribution. As a result of this, they started to report a frequency distribution for production leadtime, as can be seen in Table 2.1. This forced Westinghouse to reduce the frequency distribution for Hot and Cold Rolled Steel to a point. They accomplished this by developing a simple weighted average as shown below:

TABLE 2.1

PURCHASING MAGAZINE LEADTIME

January leadtimes this month compared to last:
Shorter: 55 Longer: 47 Same: 23

Purchasing's Leadtimes

(125 items in production quantities; % of buyers responding)

STEEL	1-5 wks	6-10 wks	11-20 wks	21-30 wks	over 30 wks
Plate	75%	19%	6%	0%	0%
Sheet & strip (HR & CR)	58	39	3	0	0
Sheet & strip (SS)	59	36	5	0	0
Galvanized sheets	62	36	2	0	0
Precoated sheets	55	36	9	0	0
Tinplate	63	25	12	0	0
Bars & rods (HR & CF)	60	33	7	0	0
Bars & rods (SS)	73	24	3	0	0
Strapping	88	8	4	0	0
Structurals	82	15	3	0	0
—Tool steel	60	20	20	0	0
Gray iron castings	18	55	27	0	0
Steel castings	7	42	47	4	0
Malleable castings	10	32	55	3	0
—Investment castings	10	42	45	3	0
—Forgings	17	31	45	7	0
Steel wire (incl. galv.)	56	44	0	0	0
Carbon tubing	55	36	9	0	0
Alloy tubing	45	37	16	2	0
Welding rods	84	13	3	0	0

NONFERROUS METALS					
Sheet & strip (copper & brass)	64	36	0	0	0
Bars & rods (copper & brass)	67	33	0	0	0
Copper tubing	55	45	0	0	0
Bronze castings	22	53	25	0	0
Bronze forgings	25	44	31	0	0
Copper wire & cable	49	41	10	0	0
Magnet wire	50	50	0	0	0
Sheet & strip (aluminum)	53	30	17	0	0
Bars & rods (aluminum)	65	30	5	0	0
Aluminum wire & cable	59	32	5	4	0
Tubing (aluminum)	57	37	6	0	0
Aluminum castings	13	64	23	0	0
Aluminum forgings	0	40	40	20	0
Bars & rods (nickel, Monel, etc.)	23	31	38	8	0
Bars & rods (titanium)	33	33	17	17	0
Die castings (all kinds)	11	48	41	0	0

FABRICATED METAL PRODUCTS					
Weldments	46	48	4	0	2
Structural steel fabricated	37	52	9	2	0
Cans	57	43	0	0	0
Steel tanks	21	53	23	3	0
Stampings, steel	26	66	8	0	0
Stampings, nonferrous	15	67	18	0	0
Jigs and fixtures	20	57	20	3	0
General machining	45	45	10	0	0
Powder metal parts	12	50	33	0	5

MATERIAL HANDLING EQUIPMENT					
Cranes & hoists	16	44	36	0	4
Lift trucks	32	25	39	4	0
Conveyors	19	38	38	0	5
Lift truck batteries	72	17	11	0	0

MACHINERY AND PARTS					
FHP motors	41	28	22	7	2
Electric motors, 1-30 hp	43	23	24	10	0
Electric motors, over 30 hp	23	39	29	9	0
Distribution transformers	19	48	33	0	0
Motor controls	39	35	24	2	0
Switchgear	17	38	33	4	8
Engines	8	31	38	23	0
Pumps, centrifugal, etc.	19	38	27	16	0
Gears	35	33	20	8	4
Nonfriction bearings	48	27	19	4	2
Screw machine parts	34	53	13	0	0
Fasteners	71	20	9	0	0
Pipe fittings	83	13	3	1	0
Machine tools	6	11	43	28	12
Repair parts (machine tool)	44	39	15	2	0
Cutting tools	72	25	3	0	0

LUMBER AND WOOD PRODUCTS	1-5 wks	6-10 wks	11-20 wks	21-30 wks	over 30 wks
Boxboards	94%	6%	0%	0%	0%
Dimension lumber	91	9	0	0	0
Pattern lumber	83	17	0	0	0
Plywood	92	5	2	1	0
Hardboard	97	3	0	0	0
Pallets	95	5	0	0	0
Wood boxes	97	3	0	0	0
Wirebound boxes	86	14	0	0	0

INSTRUMENTS AND ELECTRONIC PARTS					
Pressure gages	42	35	19	2	2
Temperature controls	40	37	21	2	0
Thermometers	50	30	20	0	0
Measuring instruments & gages	29	46	23	2	0
Chart recorders	22	48	26	4	0
Electric meters	40	45	10	5	0
Electronic test equipment	18	36	36	10	0
Switches (component)	39	41	18	2	0
Relays & solenoids	40	40	18	2	0
Transformers (component)	30	35	33	2	0
Semiconductor devices	36	39	25	0	0
Resistors, capacitors, etc.	32	36	32	0	0
Instrument motors	23	27	41	5	4
Printed circuits	19	58	23	0	0

PAPER MILL PRODUCTS					
Envelopes & stationery	92	8	0	0	0
Kraft paper	85	14	1	0	0
Printing paper	91	9	0	0	0
Bond paper	96	4	0	0	0
Offset masters	78	19	3	0	0
Tab cards	65	10	3	0	2
Corrugated containers	90	9	1	0	0
Multiwall bags	66	34	0	0	0
Foil laminates	54	42	4	0	0

GLASS AND CLAY PRODUCTS					
Glass parts	29	57	14	0	0
Glass bottles	50	39	11	0	0
Insulators	18	65	17	0	0
Refractories	17	56	27	0	0
Fiber glass	33	54	10	3	0

RUBBER AND PLASTIC PRODUCTS					
Molded rubber and plastic parts	39	56	5	0	0
Plastic film	80	16	4	0	0
Plastic laminates	55	45	0	0	0
Plastic pipe and tubing	71	21	8	0	0
Plastic bottles	69	27	4	0	0
Rubber sheeting and gaskets	67	31	2	0	0
Seals and rings	64	34	2	0	0
Conveyor belting	58	39	3	0	0
V-belts	89	9	2	0	0
Rubber hose	75	24	1	0	0

CHEMICALS					
Paint	94	6	0	0	0
Refrigerants	96	4	0	0	0
Compressed gases	97	3	0	0	0
Solvents	98	2	0	0	0
Plastic resins	75	25	0	0	0
Sulfuric acid	100	0	0	0	0
Chlorine	94	6	0	0	0
Phthalic anhydride	87	13	0	0	0
Benzene	89	11	0	0	0
Methanol	94	6	0	0	0
Anhydrous ammonia	97	3	0	0	0
Caustic soda	95	5	0	0	0
Ethylene glycol	95	5	0	0	0
Sodium carbonate	92	8	0	0	0

Table 2.2

Lead Time From *Production Magazine*

	Lead Time in Weeks				
	0-5	6-10	11-20	21-31	Over 30
Hot and Cold Rolled Steel	10%	40%	20%	15%	15%

Weighted Average = 2.5(10%)+8(40%)+15(20%)+25(15%)+30(15%)

For independent variables, 32 measures were considered most of which were economically based. A stepwise regression was run, which suggested a model using five independent variables as the best predictor for leadtime in the Hot and Cold Rolled Steel Industry. The model used is shown below:

$$y = 4.8 + 28x_1 + 9.7x2_2 + .41x3_3 + 1.6x4_4 + 4.8x5_5$$

where:

y = Production Leadtime--Hot and Cold Rolled Steel

x_1 = Capacity Utilization--Federal Reserve

x_2 = Changes in Business Inventory

x_3 = Production Leadtime Lagged one period

x_4 = Dummy for a change of leadtime index

x_5 = Dummy for a change in leadtime index

This model was tested for its predictability and was found to perform quite well (R^2 = .92).

Like most regression forecasting models, testing the past data can make the model appear to perform better than it will in actual practice. The application of the model uses predictions of future capacity utilization and inventory levels from the Data Resource Corporation. Therefore, the future success of a model relies heavily upon the success of that firm in its ability to predict those two variables.

This model was developed in 1976 and the actual implementation has not been been as extensive as originally planned. The original goal was to take predictions of PLT and load them directly into their computerized inventory control system. As this report will reveal, such a process assumes consistency among all products in terms of PLT. What the model did provide Westinghouse is a monitoring system for PLT to allow management to adjust to possible increases and decreases of overall production leadtime.

b. Aerospace Materiels Leadtime Model:

The second model that was found to be a possibility as an external forecasting model was based on a report prepared by Allen S. Davis entitled "An Investigation of Selected Business Indicators as Related to Aerospace Materials Leadtime" (7). This report was prepared in May 1974, for the Air University, Maxwell AFB, Alabama. That report notes that

during 1973-75, the leadtime of the aerospace materials and components industry increased greatly from previous normal levels, reaching a peak in the fall of 1974, and declining somewhat by February 1975. This drastic change created significant problems for DoD managers involved in the acquisition of new weapon systems and in the procurement of supporting spares for developed weapon systems. It highlighted the need for a mechanism to anticipate or predict future occurrences of increases in PLT. The study investigated the utility of using selected published economic indicators for predicting the movement production leadtimes in the aerospace materiels industry.

This study begins with an overview of the economic setting during the period of 1973-75. It then looks at production leadtime trends within the aerospace industry during that same period of time. In order to accomplish this, Davis had to establish some measure of production leadtime for the aerospace industry. In 1973, the Aerospace Industries Association established a Cost and Lead Time Quick Reaction Network within its Management Committee, the purpose of which was to provide members of that association with indications of changes in cost and PLT. The Aerospace Industrial Association collected its data by establishing seven broad category groupings of aerospace products. Within those categories, the association had a total of 54 products for which production leadtime information was provided. The model which was developed by Davis took a sample of 10 particular products from the total of 54 and used them to develop a weighted average for production leadtime in

the entire aerospace industry. This index, known as the Composite Leadtime index, was then examined in terms of a series of economic variables, in hopes that one or more of them would lead or change prior to the actual composite index (see Figure 2.1 for composite leadtime index). It should be noted that Davis chose not to use the Purchasing Magazine leadtime index since that index had been redefined during 1974 and his concern was that the change in definition might lead to some unpredictable results.

The next step in his project was to select some possible economic variables which would predict changes in leadtime for the aerospace industry. Using various criteria, Davis selected below six variables as potential predictors:

1. Vendor Performances - percentage of companies reporting slower deliveries.

2. Buying Policy - Production Materials - percentage of companies reporting commitments 60 days or longer.

3. Condition of Manufacturers' Inventory.

4. Ratio of Unfilled Orders to Shipments - Manufacturers' durable goods industry.

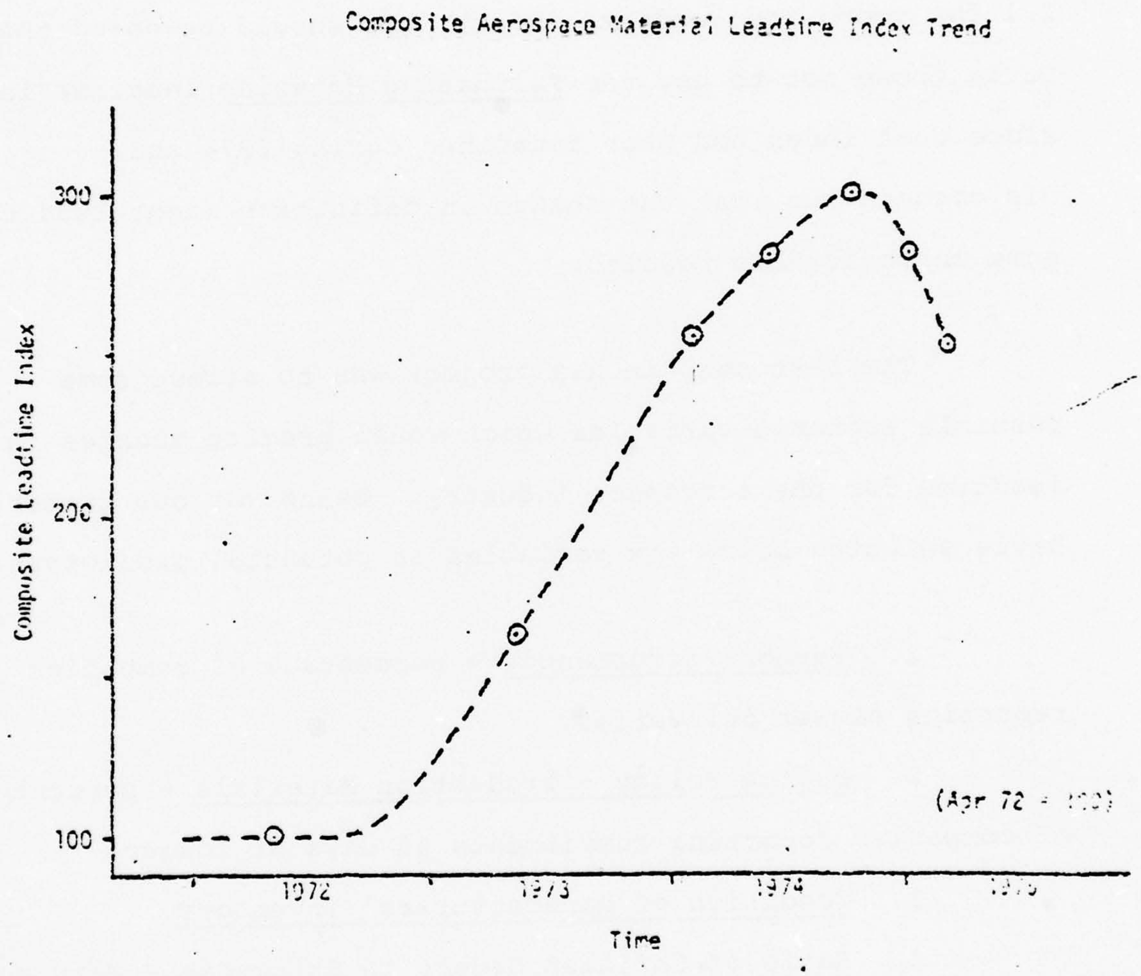
5. Adequacy of Manufacturers' Capacity.

6. Ratio of Output to Capacity - Manufacturer.

Figures 2.2, 2.3, 2.4 and 2.5 show the relationship between the composite leadtime index and the six indexes that were chosen. His method of determining whether these variables

FIGURE 2.1

COMPOSITE LEADTIME INDEX



would be good predictors of leadtime, was to simply view each of the variables along with the composite leadtime index to see whether any of the variables would have the same pattern as the index, and also change prior to that index. Figure 2.2 shows the relationship between the leadtime index, recorded quarterly, and buying policy along with unfilled orders. As can be seen by that figure, those two indices seem to parallel the leadtime index rather than lead or lag it. The next figure shows the composite leadtime index along with Book Value of Inventory. As can be seen by that figure, Book Value of Inventory changes prior to the leadtime index. The Book Value of Inventory started to move up strongly in mid-1971 almost a year prior to the movement of the composite index above 100. Book value inventory started to move down somewhere early in 1973, while the composite index didn't take its down turn until about mid-1974. He therefore concluded that Book Value of Inventory would be an excellent economic variable to use in terms of indicating whether leadtime in the aerospace industry might increase or decrease.

The next figure shows Vendor Performance Values along with the composite leadtime index. As can be seen by that figure, there again seems to be a leading indication from Vendor Performance but the amount of lead that it provides is more in the range of three or four months, rather than the nine or ten months seen in Book Value of Inventory.

FIGURE 2.2

BUYING POLICY AND UNFILLED ORDERS

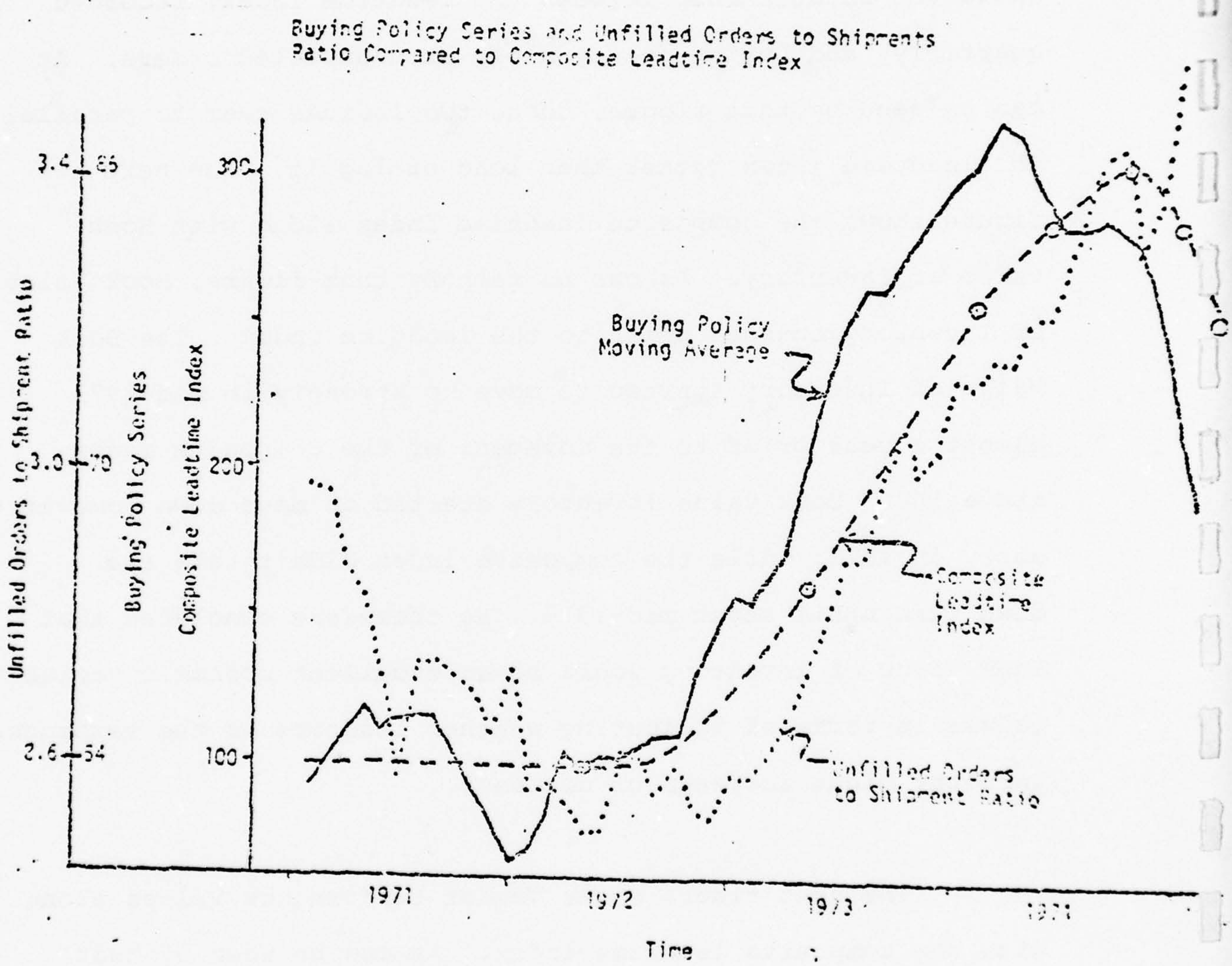


FIGURE 2.3

BOOKVALUE OF INVENTORY

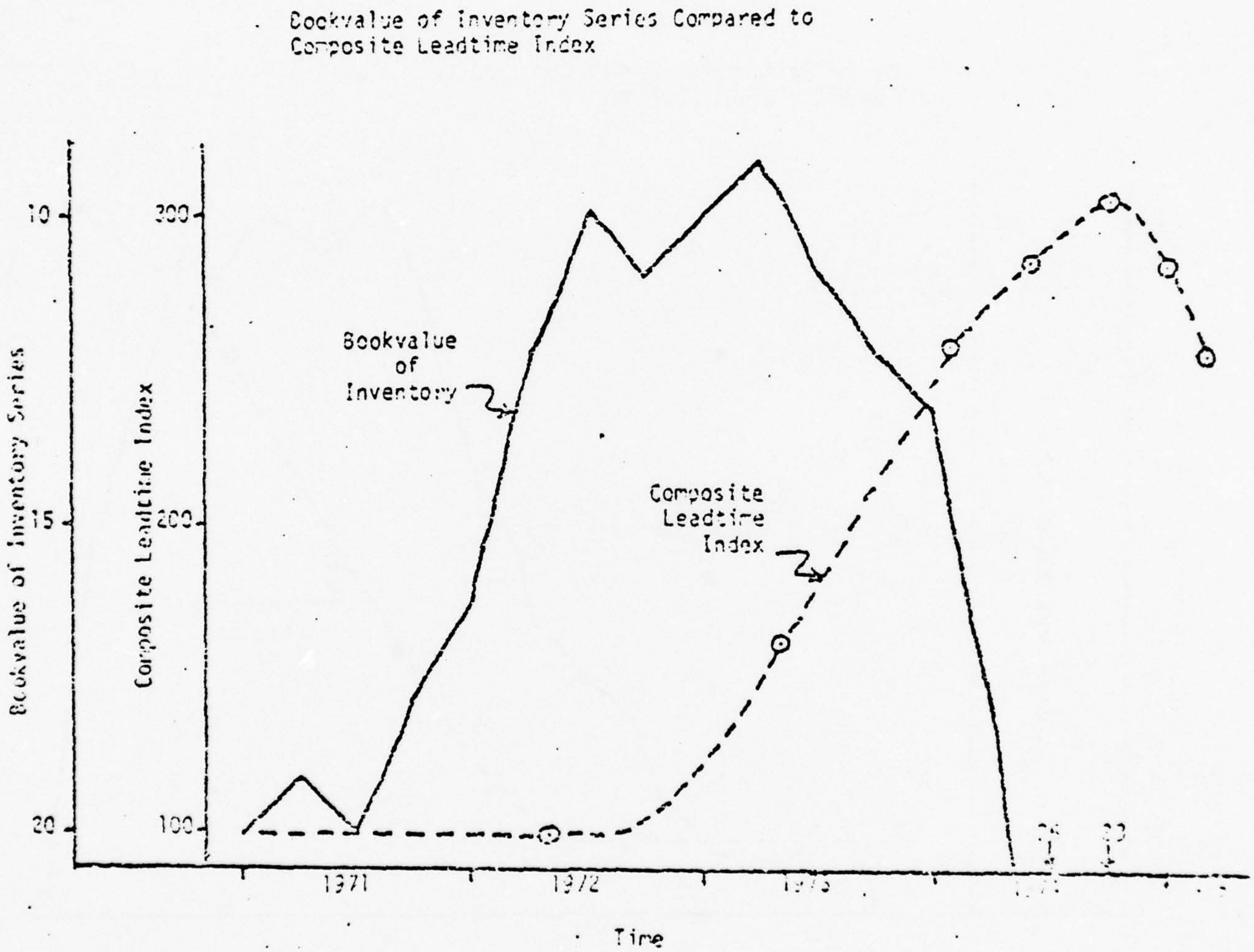
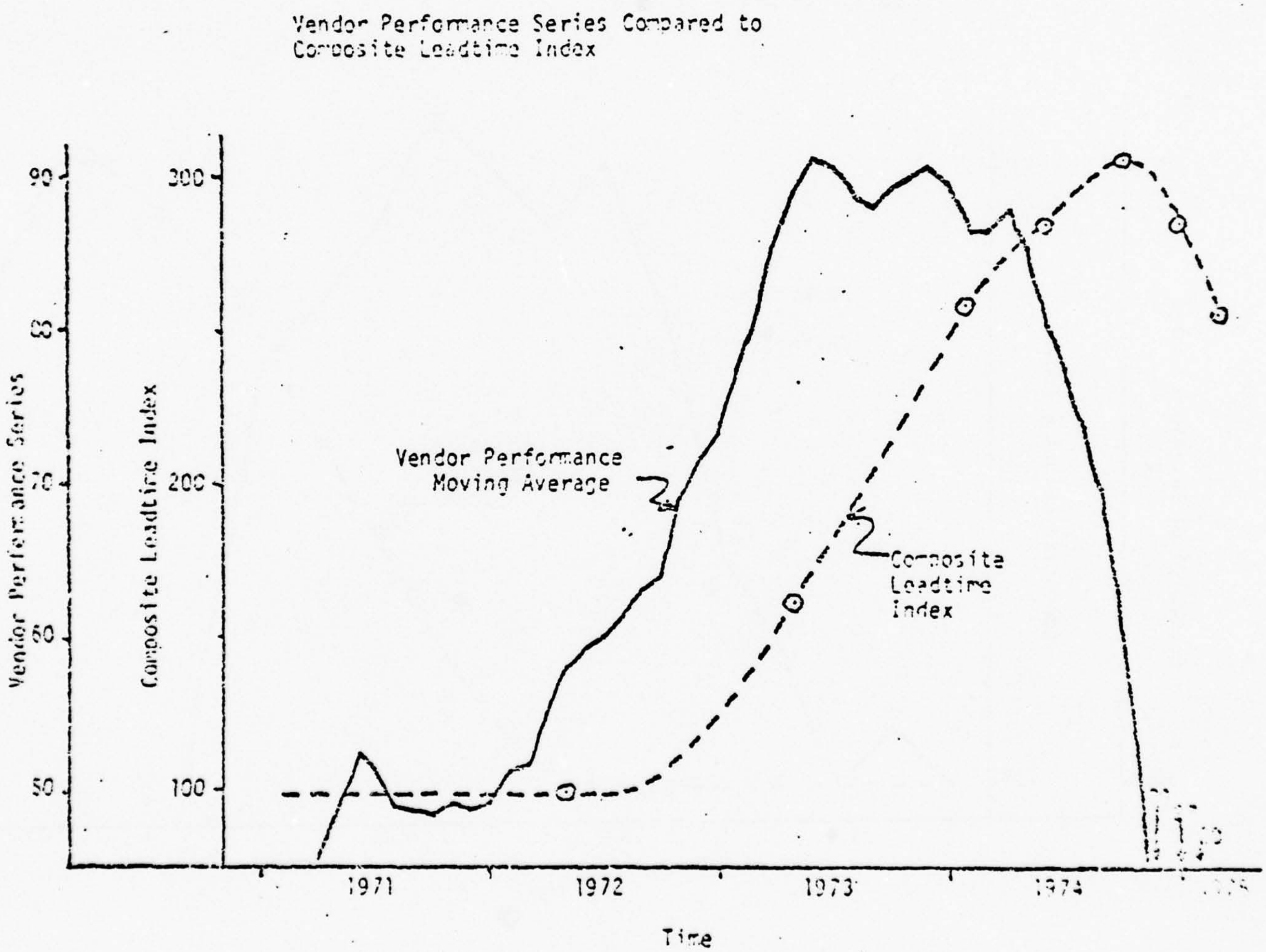


FIGURE 2.4

VENDOR PERFORMANCE



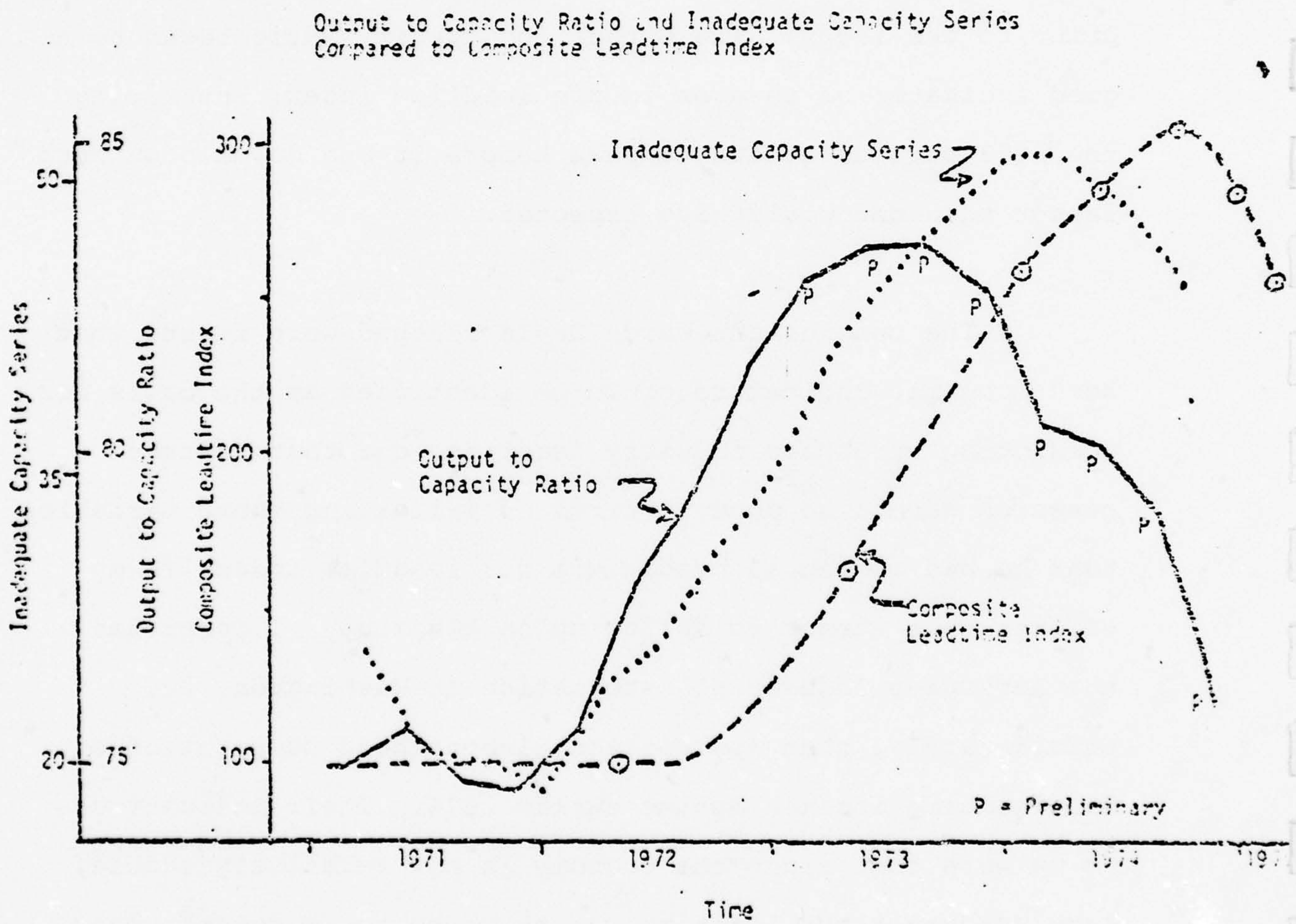
The final figure shows Output Capacity Ratio and Inadequate Capacity Series along with the composite leadtime index. The Inadequate Capacity series seems to be somewhat helpful in predicting the composite leadtime index by showing a pattern very similar to that index and consistently moving prior to the index. The Output to Capacity Ratio began as a good indicator of changes in the leadtime index, but during the 1973 period, tended to peak before it and moved down much faster than one would have expected.

The conclusions which Davis reached were first, that some economic indicators could be identified as the basis for predicting aerospace industry leadtime, and that further research should be done in terms of validating those variables that he had chosen as predictors for leadtime index. In my efforts this summer to follow up on his study, I contacted the Aerospace Industrial Association in Washington, DC. Unfortunately, that association discontinued computation of its leadtime index sometime during 1974. Their indications to me were that since the economy is now relatively stable, less concern about leadtime was shown by their constituency and therefore they felt that additional research in this area would not be meaningful to them at this time.

c. Missile Repair Model:

FIGURE 2.5

OUTPUT CAPACITY RATIO



The next model to be discussed was drawn from a report for the USAMC Interim Training Center, Texarkana, Texas, entitled "Analysis of Production Leadtime for Missile Repair Parts Contracts Dealing with Cable Assemblies and Wiring Harnesses" (10). This model was not initiated as a result of any tremendous variations in leadtime during the 73-75 period, but rather as a need to provide a better general estimating time for production leadtime. The approach is much different than in the prior models discussed, since the assumption that the author makes is that production leadtime can be related to the various characteristics of items being purchased. In essence, he is making an assumption that changes in the economic climate will have no effect upon leadtimes, but rather that leadtime is driven by the type of product being purchased.

The basic model employed examined production leadtime as it related to quantity of each buy, cost of each buy, and number of units bought during each buy. Using a stepwise regression he found total contract costs as the single largest influence on production leadtime. Further, he found that the predictability of the model was not as great as he had hoped. More particularly, the R squares were somewhere around 25 and the standard errors were great enough to warrant concern about the practicability of applying such a technique. In the conclusion, he does note that outside influence such as the economy should be integrated into a system in which one would hope to minimize the prediction error for production leadtime.

d. Distribution Model:

Another approach to this problem is to identify what the distribution of production leadtime looks like. In other words, this approach states that the first step to an understanding of production leadtime is to apply some descriptive statistics so that one can get an understanding of the mean, the variance, and the shape of that distribution. This particular approach seems to occur more on the base level than for the ALCs. Two studies were done by Aubrey Yawitz which attempted to make an evaluation of both administrative leadtime and production leadtime (27, 28). The first study examined how closely production leadtime, as estimated by the inventory system then in use, related to the actual leadtime that took place for various purchases. Generally he found that there were great differences between the estimated leadtime within the inventory system and those actually encountered. The second paper that was done was entitled "Variability of Administrative Leadtime and Production Leadtime for Troscom Managed High Velocity Items." He again looked at the distribution of a variety of items in order to gain an understanding of leadtime for various products. His design then was to relate the estimated production leadtime to actual production leadtime. The thrust of this model was to provide each item manager with a picture of the distribution of production leadtime for various items. In addition, a simple indexing method, based upon regression, would allow item managers to insert

production leadtimes for each item. The problem with his study is that it is macro-based and as indicated in some of the prior models, individual items can be influenced by a variety of exogenous factors which would make a macro-approach to this problem infeasible.

e. GERT - Production Leadtime:

Another study which attempted to look at the distribution of leadtime was done in 1973 by Kim Short (22). His model utilized the Graphic Review and Evaluation Technique (GERT) in attempting to determine distribution of leadtime. His hope was to develop a model in which one could take a category of products and talk in terms of probabilities of getting those categories on hand given some initial ordering day. In essence, he wanted to build a cumulative probability distribution so that one could look at a type item and be able to say for instance, that there would be a 90% chance of getting that product within 100 days and possibly a 95% chance of getting it within 200 days. A view of Table 2.3 which has been extracted from his report is an indication of the type of model with which he had dealt. The model did use a simulation but was based on extremely sketchy data. My view is that the model is limited since it it aggregated by types of products, whereas item managers order individual products.

TABLE 2.3

CUMULATIVE DISTRIBUTION FUNCTION

FOR CATEGORY FOUR PROCUREMENTS

X (Days)	Probability That Lead Time Is Less Than Or Equals X
66.184	0.014
92.439	0.033
118.694	0.068
144.949	0.121
171.203	0.171
197.458	0.227
223.713	0.302
249.968	0.396
276.223	0.486
302.478	0.572
328.732	0.678
354.987	0.766
381.242	0.824
407.497	0.864
433.752	0.922
460.007	0.954
486.262	0.977
512.517	0.991
538.771	0.995
565.026	0.995
591.281	0.995
617.536	0.995
643.791	0.995
670.046	0.997
696.301	0.997
722.556	1.0

f. Production Leadtime Forecasting:

The final report to be discussed, "Production Leadtime Forecasting," was done in 1972 for the Institute of Logistics Research (26). This model leads back to one of the prior studies since production leadtime is related to various characteristics of the actual product. Initiation of this study was done based on the casual empiricism of the vast differences between actual production leadtime and predicted leadtime. A regression model was developed where actual production leadtime was related to estimated production leadtime, the price of the product purchased, the quantity of the product purchased, an interaction term from the prior two items, and finally the log of an ordinal variable based on the size of the operation that the order originated from. In terms of the results, the standard errors were extremely large, and therefore the model left much to be desired.

Finally, my research included gathering information on what has become known as the Early Warning System. As an outgrowth of the 1973-75 era, the federal government decided to establish the Commodity Early Warning System. This system was a large undertaking in which approximately 150 individuals who are known as commodity specialists have the responsibility of monitoring various commodity groups. These specialists investigate items down to either the five or seven digit

Standard Industrial Classification Level. Their responsibility is to examine these industries in terms of shortages, increases in cost, changes in leadtime, or dimensions in which change in the productivity of that item would have influence for the general economy. Information from this organization is disseminate through a quarterly report which goes to any federal agency requesting such information. As of now, that report does not go to the general public, because of a fear that persons would take advantage of such information by hedging or buying futures of products which the agency predicts will face shortages.

This effort to develop the Early Warning System has gone through a series of stages and has been extremely expensive to organize. At this time there has not been an evaluation of the usefulness of the data to various agencies. As will be noted below, this type of information has so far not found its way into the AFLC, but might be extremely useful if proven to be accurate.

Summary of Survey of Literature

This chapter has looked at prior research in the area of production leadtime. In general, there are three basic avenues of attack that have been taken in attempting to identify meaningful solutions to the proper estimation of production leadtime. The first avenue has been to attempt to develop models which will predict changes in production

leadtime based upon the use of either economic variables, or variables which describe the items themselves. Those models that have tried to predict production leadtime based upon economic models have shown some success such as in the case of the Westinghouse model. One problem with that model is that the dependent variable must be one of the 125 categories that Purchasing Magazine defines. ALC does not necessarily buy primary goods, but more often than not it buys final products. These final products are defined by federal stock numbers, which have little to do with the 125 categories chosen by Purchasing Magazine.

Again, looking at the forecasting model, the second method was to try to establish characteristics of individual items and to identify whether they would be good predictors of actual production leadtime. As the survey of literature indicates, although improvements in estimates of productions leadtime have been found, the level of errors remains extremely high, and those that have developed these models admit that the lack of information about the economy and other factors can make individual forecasts very inaccurate. Thus, both forecasting models, although possibly providing methods better than presently used, have weaknesses in terms of actual application.

The second general method discussed was the process by which one simply established the form of the distribution of production leadtime for various Items. Most of these studies have focused on base operational problems. It seems that from the base view, to gain a better understanding of how vendors typically behave with various product groups, is indeed valuable information. The applicability of this kind of a model to an ALC is somewhat questionable.

The final possible solution to prediction of production leadtime is the possible utilization of the Commodity Early Warning System. This particular system came to my attention somewhat late in the 10-week program and therefore I have had less opportunity to investigate actual implementation of that type process. What we really have to do is wait to see how effective that agency is in terms of identifying areas which do indeed have increases in production leadtime. If they are capable of doing such, there might be opportunities to gather information and forward it to item managers in certain areas. The only problem will be that ALC purchases so many products that the actual implementation of such a process might become extremely difficult.

CHAPTER III

Methods of Solution

Introduction: As indicated in Chapter I, there are two possible approaches for dealing with more accurate prediction of leadtime: First, development of a model which externally determines production leadtime forecasts based on information outside of the AFLC data systems; secondly, taking the information internal to the data systems and developing some sort of smoothing or averaging techniques to predict production leadtime. The two approaches differ not only in how they would predict, but also in terms of how they would actually be implemented into the system. I expended the vast majority of my efforts on the external model for a variety of reasons. First, Procurement had done substantial research in this area yet they had not completed their model. They were therefore eager to know if their basic model had any promise for actual application. The second reason was that the Westinghouse model mentioned in Chapter II, seemed to be more of a final product than many of the other models discussed in the survey of literature. Finally, the data sources needed for some sort of experiment were readily available for this type of model. On the other hand, there were a series of negatives associated with the internal system. First was the inability

to get leadtime information out of either the D041 or D062 during the ten-week period given for this study. The second was the usefulness of that data in its present form. More specifically, the D062 system includes the update of production leadtime from the last buy, as an estimate of future demand. Presently, that information is inputted by hand through the item manager and there is no way of telling how diligently this task is performed. Therefore, the historical information in the file are estimates of production lead-time, which could have nothing to do with actual lead time experiences.

External Model:

a. Model: The basic model designed was a regression, with production leadtime as a dependent variable and a series of economic variables as predictors (11).

$$y = a + b_1x_1 + b_2x_2 + b_3x_3 + b_4x_4 + b_5x_5$$

where:

y = Production Lead Time for materiels handling equipment industry.

(a) cranes and hoists

(b) lift trucks

(c) conveyors

(d) lift truck batteries

x_1 = Industrial Production Index (defense and equipment)

x_2 = Industrial Production Index (transportation industry)

x_3 = Book Value of Inventory (machinery, except electrical)

x_4 = Unfilled Orders (machinery, except electrical)

x_5 = Wholesale Price Index (machinery,and equipment)

Although there seem to be reasons to lag some of these variables, the CREATE systems version of the SPSS did not have this capability, so our hope was to find that some of the independent variables would lead production leadtime (7).

b. Data

The first problem was choosing a dependent variable. As can be seen by Table 2.1 the 125 items that are followed by Purchasing Magazine vary from final products, such as switches or electronic meters to raw materials such as galvanized steel sheets or alloy tubing. Since the Air

Force typically purchases final products, I wanted to choose an area of finished goods versus raw materials. In addition I wanted to choose a product(s) that might be able to be identified within the D041 or D062 inventory system. As noted before, the 125 items have not been chosen on the basis of standard industrial classification (SIC) nor federal stock numbers (FSN), and therfore the only way to choose an item(s) was by compromise. I identified an individual in Procurement who was in charge of cataloging items by FSN. I discovered that many of the items in Purchasing Magazine had to be eliminated due to the uniqueness of what the Air Force actually buys. Finally, I chose the Material Handling Equipment Industry because the Air Force buys these units in their final form.

The second issue was to choose independent variables. One criterion was that the data used be easily accessible so that actual implementation of the model would not be difficult. The second criterion was that the data be recorded on a monthly basis. My concern was that the dependent variable was recorded monthly and therefore its conversion to quarterly data would result in a loss of information. In addition, procurement seemed to feel that one must stay abreast of production leadtime, and in a quarter, conditions might change enough that problems in the procurement process could arise.

Given these criteria I chose five independent variables, the first of which was an Industrial Production Index for the Defense Industry. My hope for this variable was that the activity level for this industry might have some effect upon production leadtime for the materials handling equipment industry. The second independent variable, Industrial Production for the Transportation Industry, was chosen because it should be directly associate with dependent variable of production leadtime for the materials handling equipment industry. The third independent variable was book value of inventory for machinery except electrical. This variable was chosen, based on the Davis model, which identified inventory as a good lead predictor for the aerospace industry. Unfilled orders were my fourth independent variable and the choice was based upon the assumption that it might act as a surrogate for increases or decreases in production leadtime. My final independent variable was the wholesale price index for machinery and equipment. It is well known that when demand for the product increases, not only does the production leadtime increase, but also the price of the product typically increases. This could be caused partially by the increased cost of dealing with the increased demand such as overtime, increase breakdowns, etc. along with the typical market reaction of industry to attempt to increase prices of the product in times of decreasing demand.

c. Procedures.

A data file was built on the CREATE system for lead time which went back to 1969 and covered the period up to March 1978 which reflects the latest published data on my independent variables. This data file was used as input to both an SSS and ONNITAB file.

Internal Model: As noted above only marginal consideration was given to the development of an internal model. The main reason for this was the unavailability of data from the D041 and D062 system. Yet it seemed fruitful, given the extensive literature search, to at least look at some possible uses of smoothing techniques in that it might be possible to eventually get data to test such models.

It should be recalled that the system we are about to implement in the D062 will automatically put the production leadtime of the most current buy as the estimate of future leadtime. If one assumes that production leadtime is a stationary rather than a variable, then the sample of one is a valid indication of what the production leadtime will be in the future. If the item manager has no information other than what is in the system, then some sort of smoothing

technique would have to be an improvement over the last buy criteria. It should also be noted, that if the random variable moves overtime, that the external model might be used to adjust the estimate.

For the purposes of this study let us assume that the leadtime data in the lift truck industry, as recorded by Purchasing Magazine, was fairly reflective of what was in the D041 for fork lifts. As can be seen by Figure 4.2 the leadtime started to go up in late 1973, peaked in 1975, then went down to form a steady pattern which continues through the current time period. Given a series like this, there are a variety of smoothing techniques that seem to pose promise for integrating into both a D041 and D062 system.*

a. Last Period Demand

$$\bar{Y}_t = Y_{t-1}$$

Where

$\bar{Y}_t$ = Forecast Demand for Period t

Y_{t-1} = Actual Demand In the Previous period.

* NOTE: This section simply reviews some smoothing techniques, and should be skipped if the reader is familar with them.

This is the most simple model, and it assumes that demand in any future period will be exactly the same as the present period. Such a model can be a bad predictor if the data is subject to large random fluxuation, or has seasonality. On the other hand if there is a strong trend effect, along with a small standard deviation, then it can be very effective.

b. Arithmetic Average

$$\overline{Y}_t = \frac{1}{m} \sum_{i=1}^{m} Y_i$$

(1) The basic source for this section was Tersim, Materials Management and Inventory Systems. (1977)

where:

$\overline{Y}_t$ = Forecast Demand for period t,

Y_i = Actual Demand in period i

m = Number of time periods

This model will be effective if demand is not subject to trend changes, or seasonal variations. In essence, it smooths out variations.

c. Moving Averages

$$\overline{Y}_t = \frac{1}{m} \sum_{i=1}^{m} Y_{t-i}$$

Where

$\overline{Y}_t$ = Forecast Demand for period t

Y_{t-i} = Actual demand in period t - i

m = Number of time periods included in moving averages.

This model will be good for trend changes, but will lag the trend. The amount of the lag will be a function of the number of period chooses for the moving average. Seasonal variations are not dealt with in using this model.

d. Regression Analysis

$$\overline{Y} = \alpha + \beta t$$

where

$$\beta = \frac{n \sum_{i=1}^{n} t_i Y_i - \sum_{i=1}^{n} t_i \sum_{i=1}^{n} Y_i}{n \sum_{i=1}^{n} t_i^2 - \left(\sum_{i=1}^{n} t_i \right)^2}$$

$$\alpha = \overline{Y} - \beta t$$

n = number of periods covered

t = time period

Regression is excellent when the data has trend effects. The above model shows simple linear regression, but non-lineal situations cases can also be dealt with. The model does not compensate for seasonality, and will be ineffective with respect to this variable. The variations in demands follow a pattern over time, know as autocorrelation.

e. Exponentically Weighted Moving Average

$$\bar{Y}_t = aY_{t-1} + a(1-a)Y_{t-2} + a(1-a)^2 Y_{t-3} \ldots\ldots\ldots\ldots + a(1-a)^{m+1} Y_{t-m} + (1-a)^{m+2} Y_{t-m}$$

Where

$\bar{Y}_t$ = Estimate of Demand in period t

Y_{t-1} = Actual Demand in t - 1

a = Exponential smoothing constant and between 0 and 1

The above model uses all prior data in any estimation of future demand. The exponential smoothing constant controls the importance of present verses current demands. The closer a is to 1, the more important current information, and vice versa.

The advantages to expontial smoothing are the control of importance of current versus past data, along with small computer file storage needed for implementation of the model.

f. Exponential Weighted Moving Average with Trend Correction

$$\bar{Y}_t = aY_{t-1} + (1-a)(\bar{Y}_{t-1} + T_{t-1})$$

Where

$$T_t = b\,(\bar{Y}_t - \bar{Y}_{t-1}) + (1-b)\,T_{t-1}$$

b = Exponential smoothing constand (0 to 1) to adjust for the trend

This model will integrate a trend correction into the smoothing process. The choice of b is again dictated on the desire to have greater weight for current versus past time periods.

g. Exponential Weighted Moving Average with Seasonal Correction

$$\bar{Y}_t = \left[a\,(Y_{t-1}) + (1-a)\,\bar{Y}_{t-1}\right] \frac{I_t}{I_{t-1}}$$

Where:

$$I_{t+m} = c\,I\left(\frac{Y_t}{\bar{Y}_t}\right) + (1+c)\,I_t$$

I_t = Seasonal Index for period t

c = Exponential smoothing constant between 0 and 1

m = number of periods in season

This model will adjust for seasonal varations in the data. The choice of c will dicatate the importance of past versus current seasons variations.

h. Exponential Weighted Moving Averages with both Seasonal and Trend Corrections

$$\bar{Y}_t = \left[a Y_{t-1} + (1-a)\left(\bar{Y}_{t-1} + T_{t-1}\right)\right] \frac{I_t}{I_{t-1}}$$

When both seasonal and trend are included, then the trend calculation has to be adjusted.

$$T_t = b\left[\frac{\bar{Y}_t}{I_t} - \frac{\bar{Y}_{t-1}}{I_{t-1}}\right] + (1-b)\,T_{t-1}$$

i. In reviewing the above forecasting techniques it seems that exponention smoothing with trend would be the best clearly has significant trend effects in it. In addition I tested the data for seasonality, and found that it did not exist.

CHAPTER IV
MODEL RESULTS

As indicated in Chapter III, the only model to be implemented was the external model. Therefore this chapter will briefly examine the results of the regression model developed in that chapter. The procedure will be to notice how each of the independent variables look over the nine year period and then view the independent variables over the same time interval. Finally, we can observe the nature of the regression model in terms of its statistical significance and accuracy.

DEPENDENT VARIABLE

Figures 4.1 through 4.5 show the four independent variables, plus a composite, for the period of January 1969 through July 1978. The pattern in all four cases is similar. That is a rise in production leadtime takes place from January 1969 untill sometime in the beginning of 1970, and then takes a downturn untill January of 1973. From this point on, leadtime takes a bold upsurge, which lasts until late 1974. A pronounced peak is then formed, with leadtimes falling as dramatically as they had been so that late 1975, leadtimes were back to the somewhat normal levels from prior periods. From that point to the present they have varied, but no large upsurge or downsurge has appeared.

FIGURE 4.1
CRANES & HOISTS

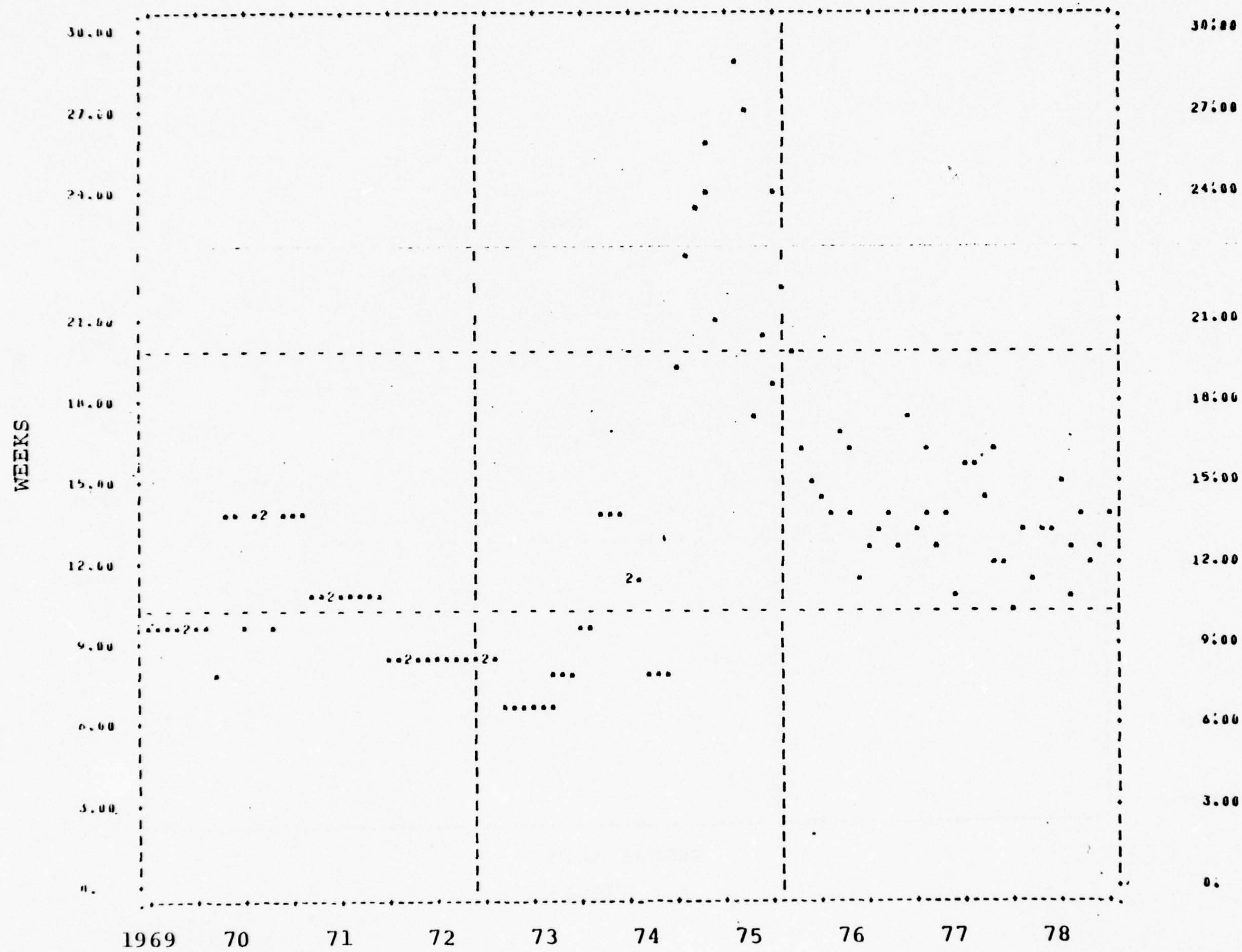


FIGURE 4.2
LIFT TRUCKS

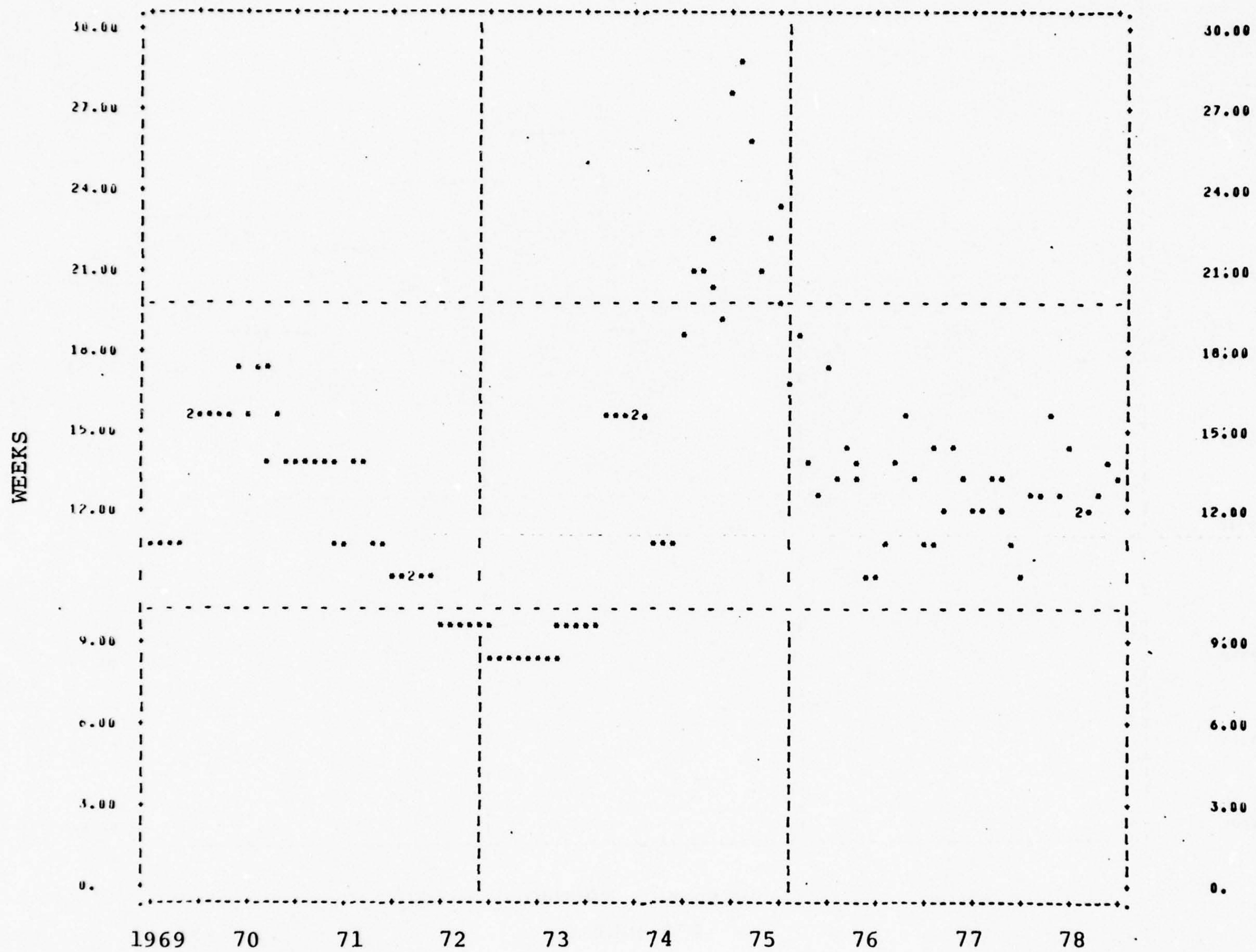


FIGURE 4.3

CONVEYORS

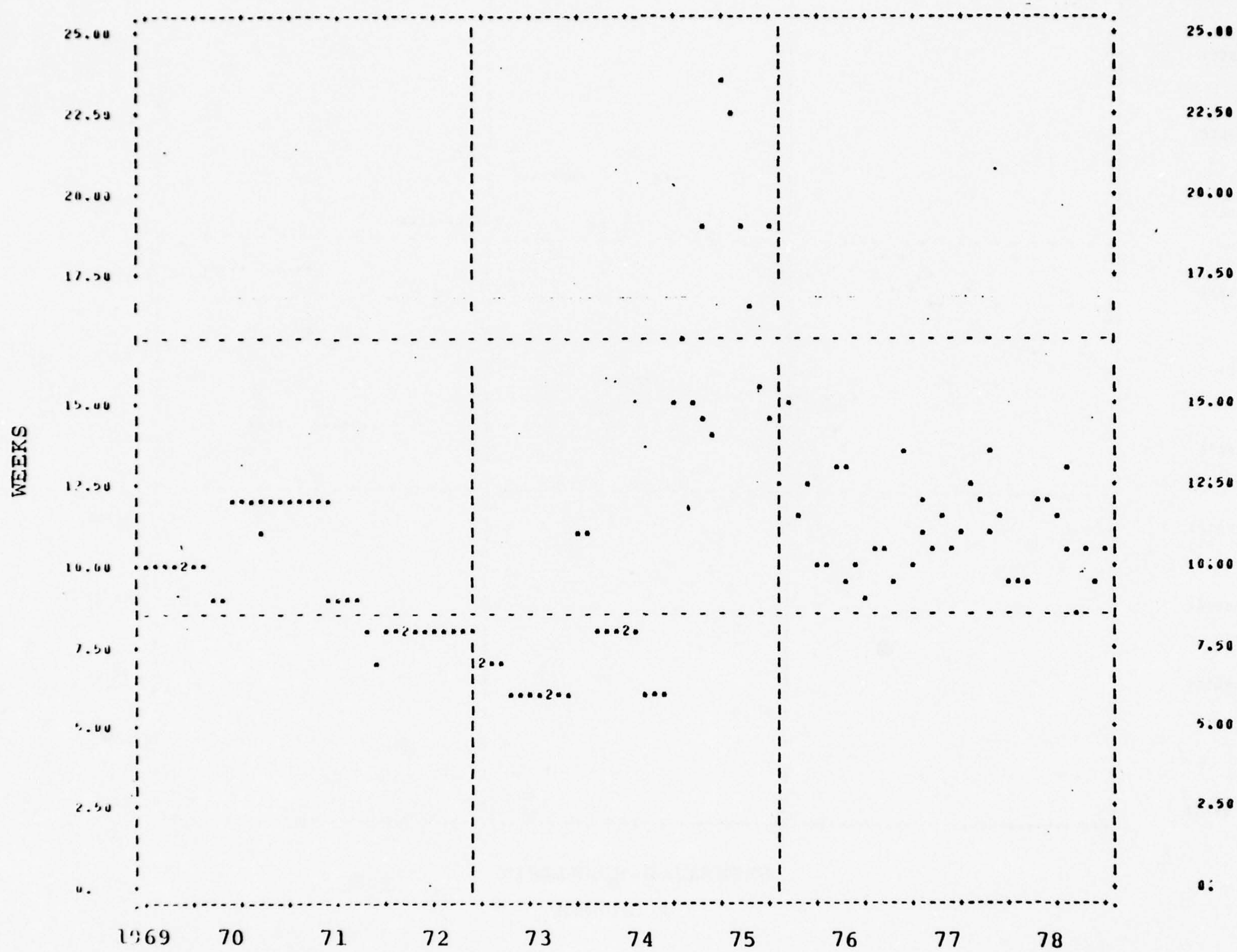


FIGURE 4.4

LIFTTRUCK BATTERIES

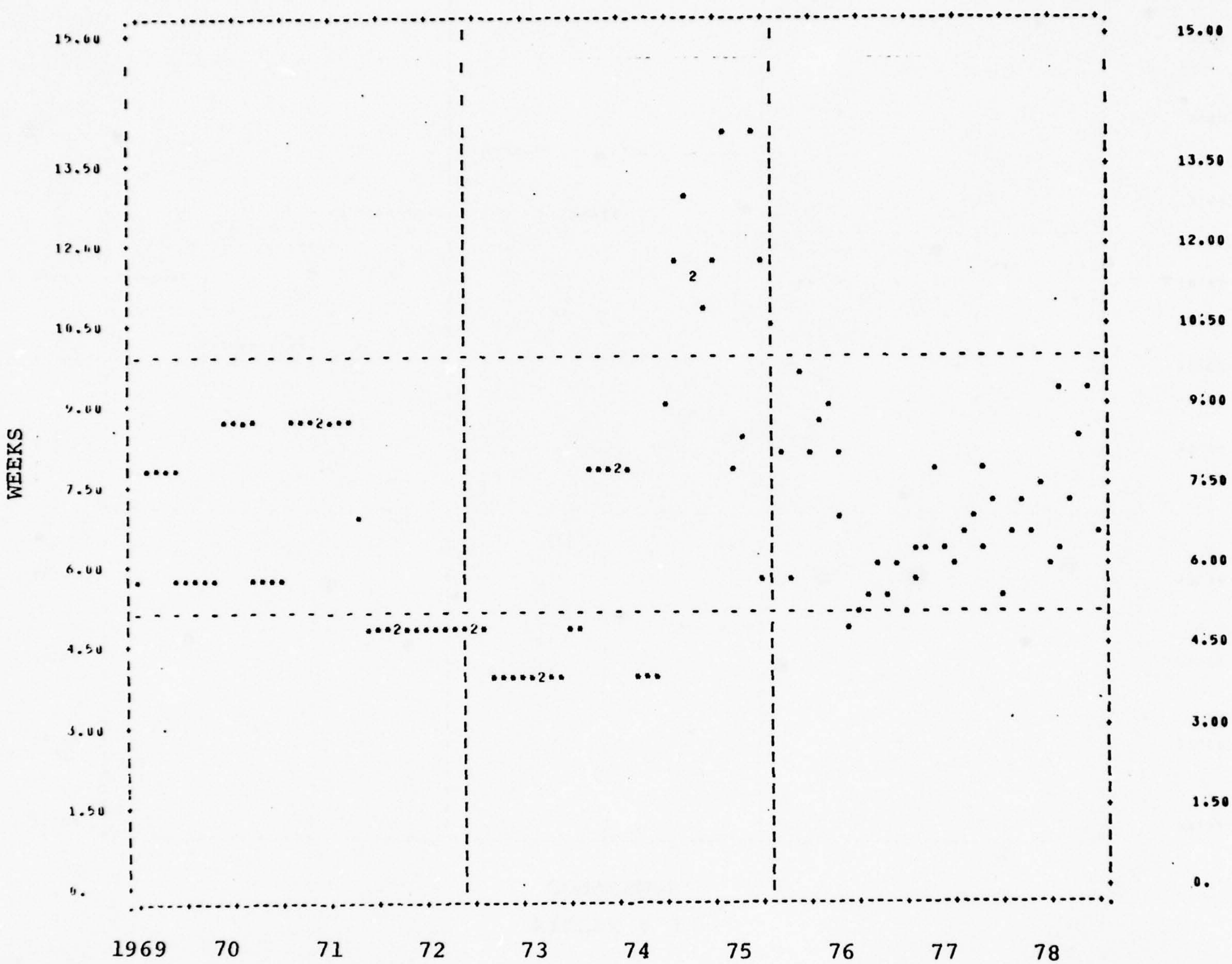
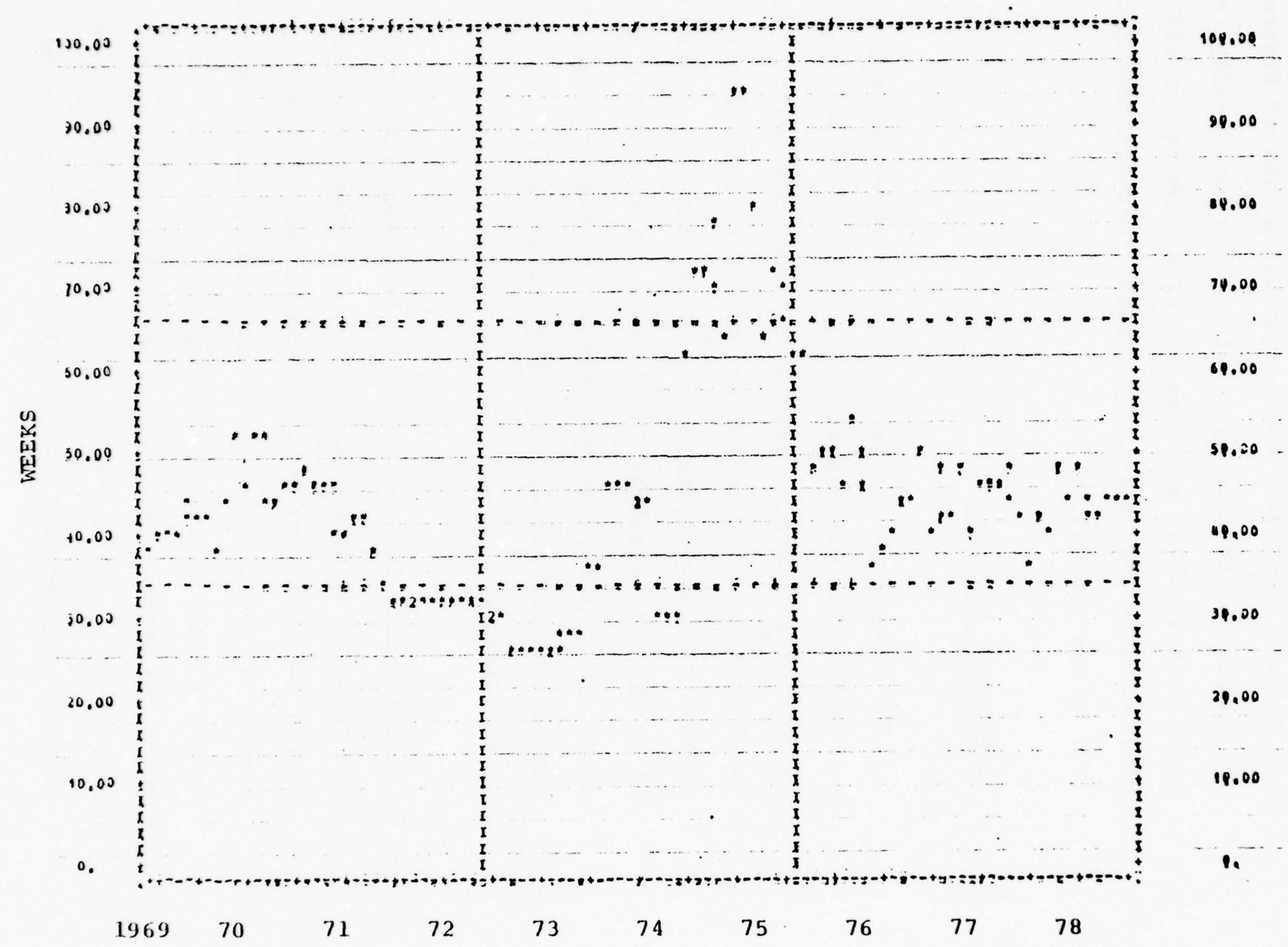


FIGURE 4.5

COMPOSITE INDEX



To best understand these changes, a view of figure 4.5 of the composite leadtime is helpful. One can clearly see the small peak in 1970, and then the great upsurge in 1974. Also observable is the relatively constant pattern from January 1976 to the present.

INDEPENDENT VARIABLE

The independent variables did not show the same consistency. (See Figures 4.6 to 4.10). The first, industrial production for defense equipment, drops off heavily in 1969 until it hits a low in January of 1971. This is the result Vietnam military deescalation. From that point until early 1974, there is a slight increase, which then falls off and remains relatively constant untill the present time. By visual inspection this variable seems quite unrelated to the dependent variables.

The second independent variable, industrial production for the transportation index, shows a different pattern. It begins similarly, by falling from January of 1969 until January of 1971, but then it increases until January of 1974 taking a violant drop in January of 1975. From that point until the present it has been showing a steep but steady increase. Visual inspection shows this variable as a possible leading indicator for 1969 to 1975, but from that point there

FIGURE 4.6
INDUSTRIAL PRODUCTION INDEX
(DEFENSE EQUIPMENT)

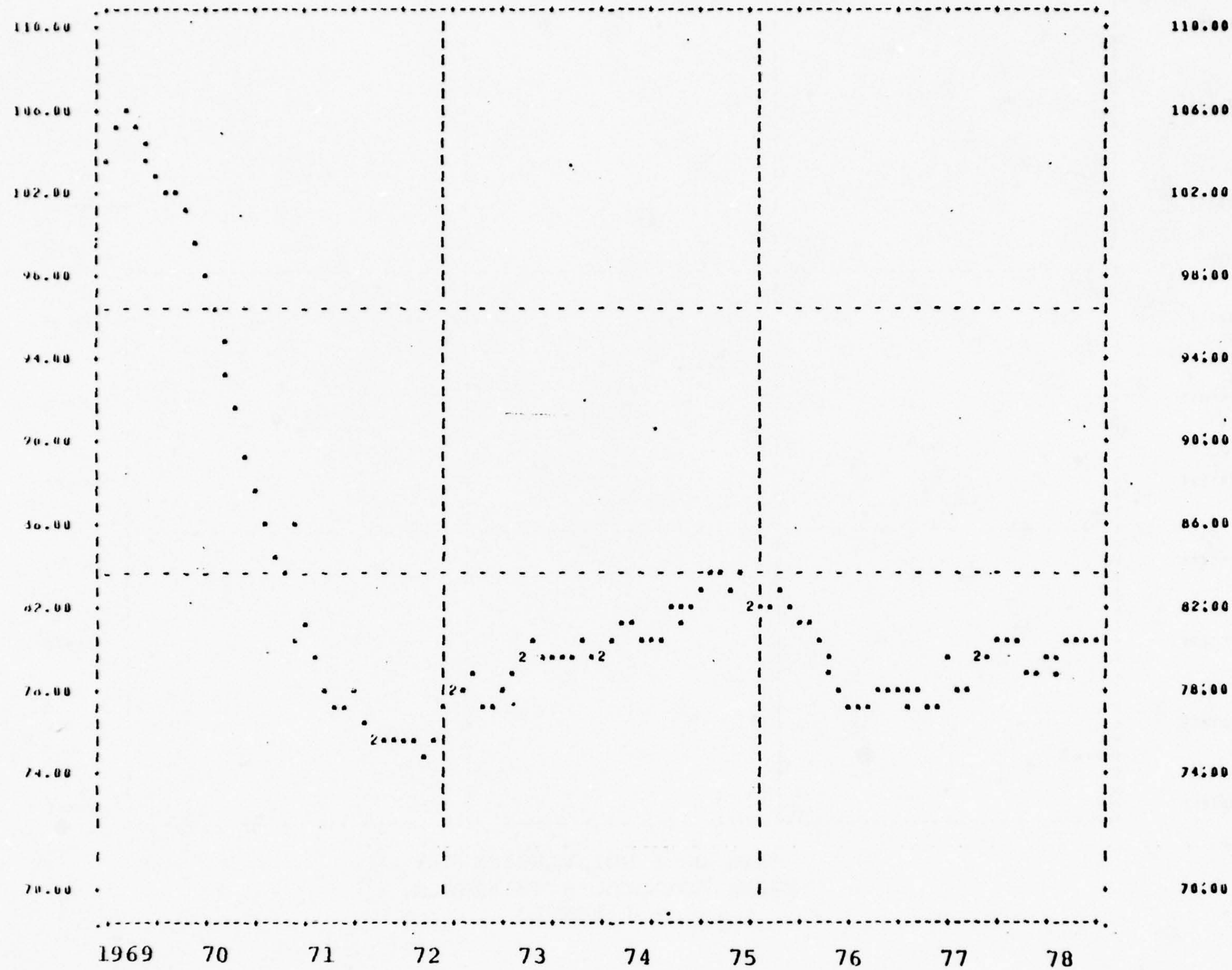
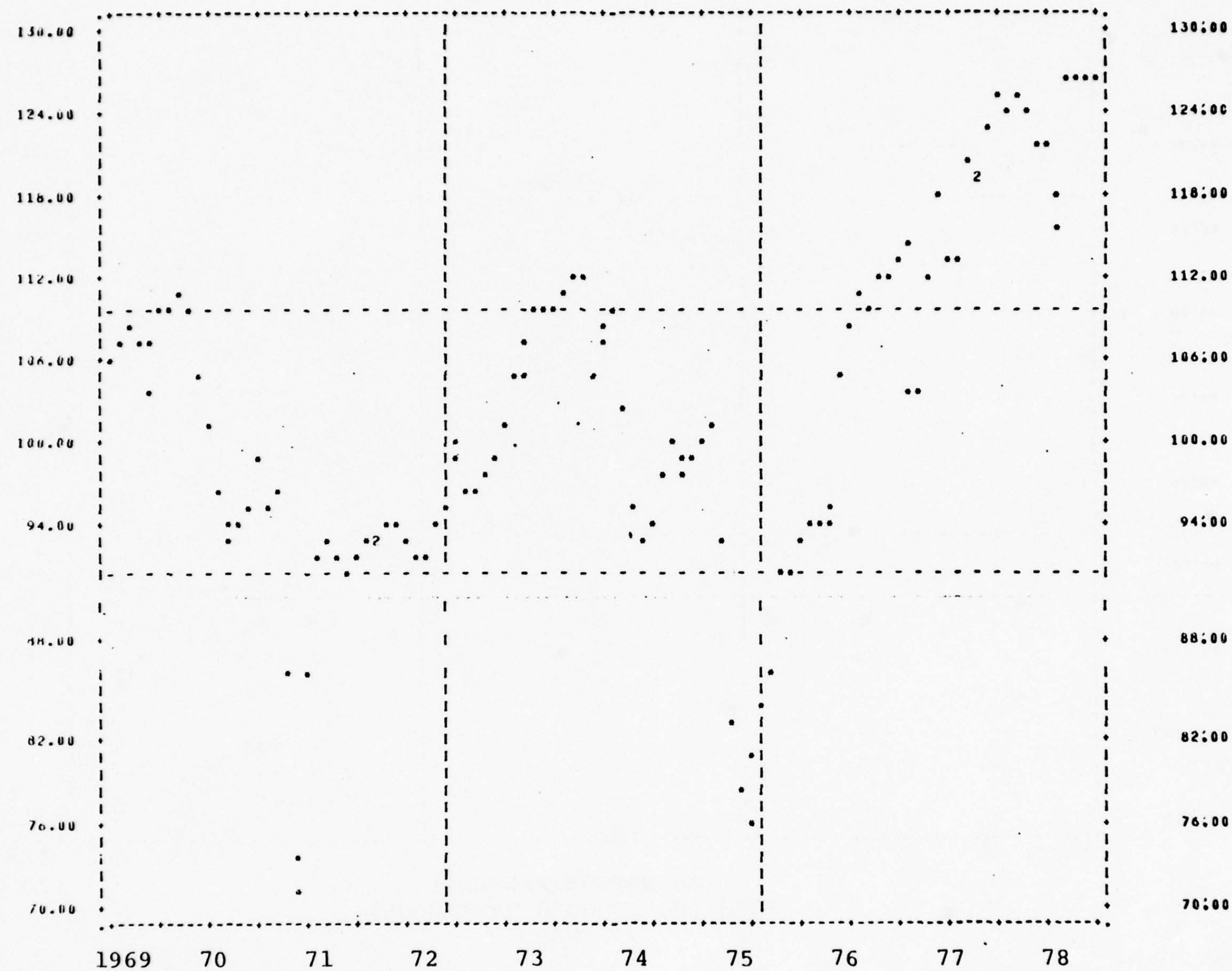


FIGURE 4.7
INDUSTRIAL PRODUCTION INDEX
(TRANSPORTATION INDUSTRY)



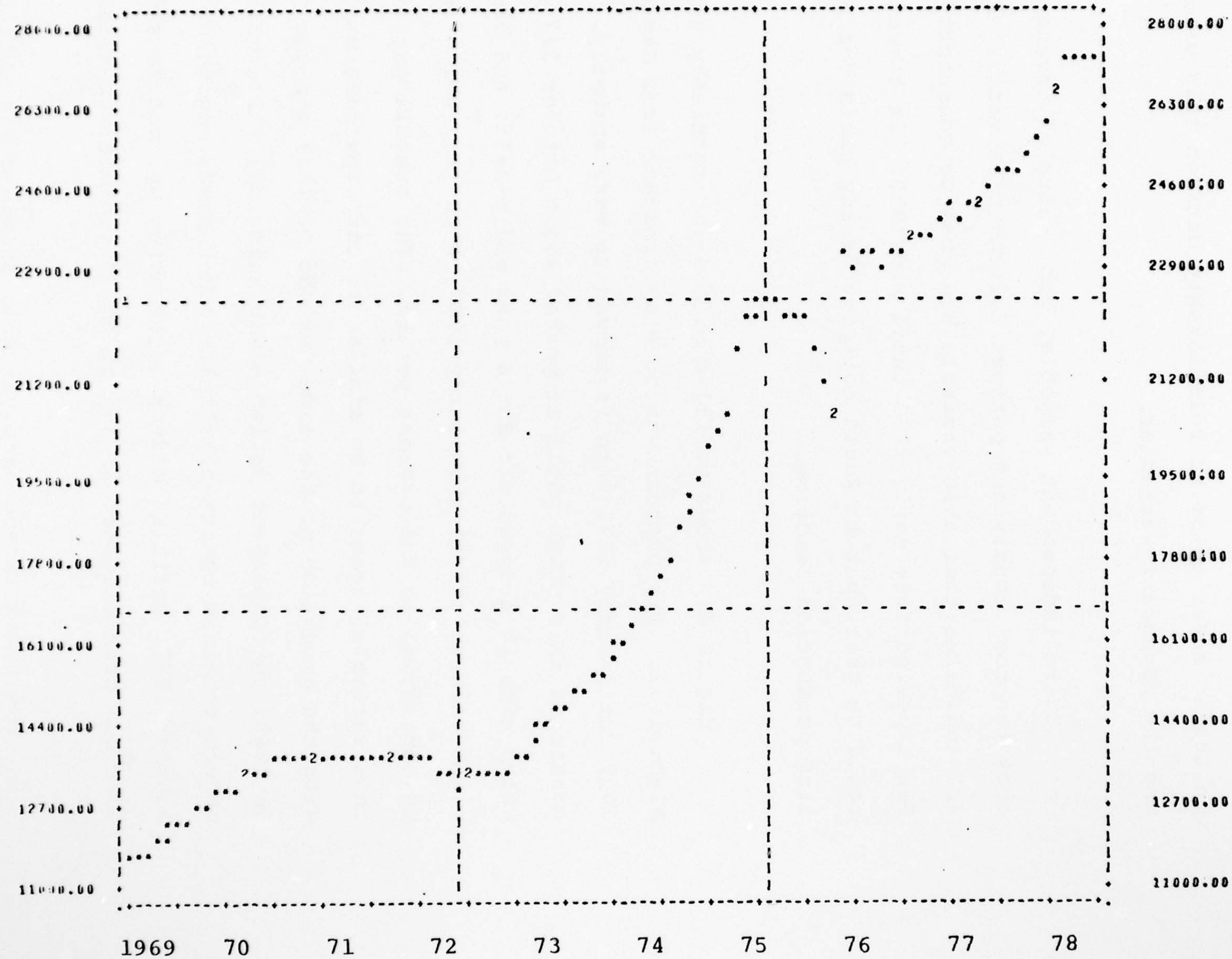
on there seems to be no relationship between this variable and the dependent variables.

Third independent variable, book value of inventory, show the most consistent pattern for the nine years. Figure 4.7 indicates that this variable has gone up constantly, except for a period from early 1970 until mid 1972. It therefore could be seen that it would not provide any basis for predicting production leadtime.

Figure 4.8 shows unfilled orders for machinery except electrical. Its pattern was to stay constant from January 1969 until early 1972 when it showed up very abruptly. It continued this climb until it peaked about October 1974, at which time it dropped off for a year and a half, and began a climb back up starting in early 1977, a pattern which seems to be continuing to the present period. The overall pattern of this variable seems to be similar to the dependent variables, with the exception of the time periods of 1969 and the present. In 1969, the dependent variable was going up, while unfilled orders remained constant. On the other hand, beginning in January 1977, unfilled orders began going up, but no similar turn has been detected in any of the dependent variables.

The final independent variable examined was that of the wholesale price index for machinery, except electrical.

FIGURE 4.8
BOOK VALUE OF INVENTORY



index seems to show little resemblance to any of the dependent variables. Figure 4.10 shows a steady and constant increase since 1969, with one upsurge in 1974.

As indicated in Table 4.1, each model has significant variables and fairly high R squares. Such a result is not uncommon in terms of time series data. The one variable that entered first when these models were run on stepwise was X_4, unfilled orders. This was not surprising, given the prior view of the scatter diagrams for the variables.

The usefullness of this model as a predictive device for forecasting PLT seems limited. High R^2 and F values are not substitute for the pure logic of viewing the scatter diagrams, all of which indicate that the independent variables selected are poor predictors of PLT. Because of high correlations in the independent variables, one could pursue a factor analysis, then use variables with high readings as surrogates for factors, or factor scores as the new predictive variables. Such a line of procedures might produce higher F values, but it would not change the basic problem of the model's inability to lead PLT (9). Finally, the Durbin-Watson statistic was significant for our models, indicating the presence of autocorrelation.

FIGURE 4.9
UNFILLED ORDERS

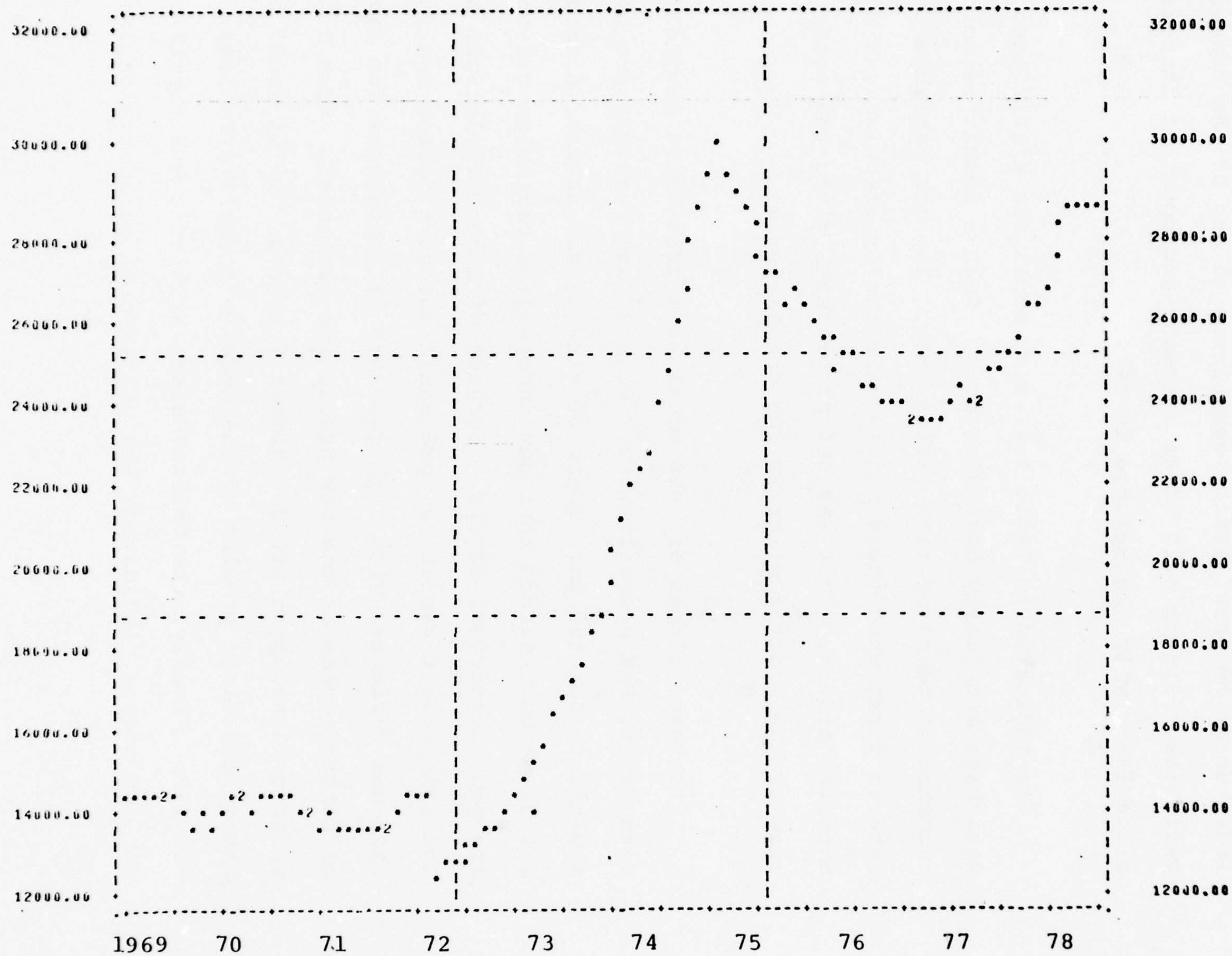


FIGURE 4.10
WHOLESALE PRICE INDEX

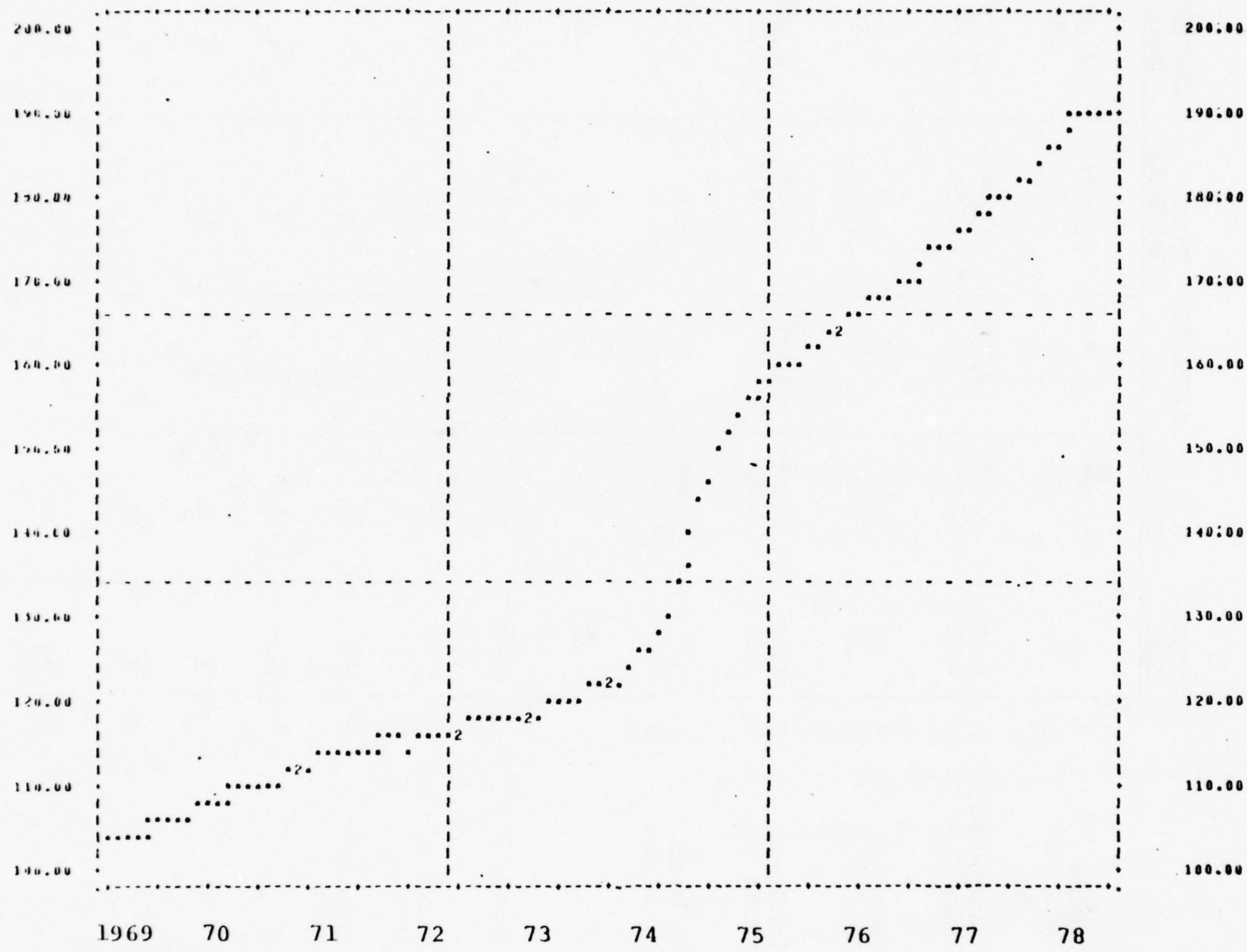


TABLE 4.1

REGRESSION RESULTS

		CRANES & HOIST	LIFT TRUCKS	CONVEYORS	TRUCK BATTERIES	COMPOSITE INDEX
Defense Index (X_1)	F	2.9	20.6**	20.7**	11.98**	16.2**
	b	.11	.29	.34	.29	.25
Transportation Index (X_2)	F	7.8**	5.31*	14.75**	7.24**	11.12**
	b	-.20	-.16	.32	-.25	-.24
Inventory Level (X_3)	F	.07	3.16	.27	1.76	1.14
	b	.17	1.13	.38	1.11	.67
Unfilled Orders (X_4)	F	50.0**	48.03**	18.9**	11.34**	44.16**
	b	1.21	1.16	.84	.74	1.11
Wholesale Price (X_5)	F	1.5	10.98**	.54	3.42	4.20*
	b	-.68	-1.79	-.46	-1.32	-1.11
R^2		.67	.69	.58	.46	.68
Overall F		44.88**	48.17**	30.43**	18.48**	47.72**
Standard Error		2.85	2.27	2.24	1.72	7.75
Order of Entry	1st	X_4	X_4	X_4	X_4	X_4
	2nd	X_5	X_5	X_5	X_5	X_5
	3rd	X_2	X_1	X_1	X_1	X_1
	4th	X_1	X_2	X_2	X_2	X_2
	5th	X_3	X_3	X_3	X_3	X_3

Note: * Significant at 5% level
** Significant at 1% level
b= Standardized Regression Coefficient

CHAPTER V

SUMMARY AND CONCLUSIONS

SUMMARY

This report has dealt with the problem of predicting Production Leadtime (PLT) for AFLC. The concern about this problem originated with the substantial increases in production leadtime for both military and industry products during the economic disorders of 1973-75. The primary reason that this dramatic change in PLT had such a profound effect upon AFLC was the method used to estimate future PLT. In essence, future production leadtime was (and still is) estimated to be the same as the actual PLT of the last buy. On surface examination, this might appear to be a reasonable approach, since as leadtime increases or deacreases, the system should simply lag one period behind. The problem is, however, that changing the production leadtime to the last buy must be done by hand by the item manager. While hard evidence is not available on the effectiveness of this system, it seems unrealistic to assume that each time an item is purchased, an item manager will in fact update the PLT. This system resulted in buy orders being initiated too late for timely deliveries.

As a result of this experience, Procurement began to develop a procedure that would anticipate changes in PLT. The first facet of that procedure was to survey contractors of

items that had not been recently purchased, in order to glean acurate predictions of current PLT. This system has been in operation for three years, but has not been investigated in terms of its effectiveness. The secont step was to periodically track current PLT in 12 industrial groups. The index used is one published in Purchasing Magazine, and is the result of a bimonthly survey of purchasing agents. This seems to provide an effective method of keeping AFLC Headquarters informed about changes in PLT. The final step taken by Procurement was the development of a forecasting model intended to predict changes in the 12 industrial groups. The basic structure of this model would be a regression, with a series of economic variables as predictors.

A survey of the literature revealed a variety of studies on leadtime done by all three brances of the armed services. The majority of the studies focused on administrative leadtime, since the services have control of this aspect of leadtime and could therefore make adjustments in procedures if new ways of reducing leadtime were revealed. In the area of production leadtime the research is sketchy and diverse. Most efforts have been directed toward establishing the shape of the distribution of PLT for individual items or groups of similar types of items. This approach has been especially attractive to those examining PLT at the base level. The logic behind this

approach is that the identification of the mean leadtime and information about variations in PLT would be a great improvement over the present systems, which often assign the same value for PLT for all items. In addition, the computation of the single PLT value is often arbitrary, and frequently not based on actual PLT experience.

The literature search did reveal two studies which attempted to forecast PLT. The first of these was done by Westinghouse, whose model attempted to predict changes in PLT for the hot and cold rolled steel industry. The method used was regression, with PLT as a dependent variable, and a series of economic variables as predictors. Although the model proved to be a good predictor, Westinghouse never completely implemented it into its inventory system because of problems in disaggregation.

The second model was developed by Allan Davis for Air Force University. This was less sophisticated, but did examine the aerospace industry. His model viewed a series of economic variables in hopes of finding one or more which would lead changes in PLT for the aerospace industry.

The one type of model not fully investigated was an internal predictive model, which would smoothe the change in

actual item PLT instead of trying to predict these changes by the use of exogenous variables.

The model developed for this program was external, along the lines of the Westinghouse model. The dependent variable was PLT for the materials handling industry. Five independent variables were chosen, all of which were recorded monthly. The time period used is from January of 1969 to the present. Of the five variables selected, only that of unfilled orders gave promise of being somewhat helpful in forecasting PLT. Although some of the other variables had significant F values, an examination of the scatter diagrams indicate that they would be poor predictors of production leadtime.

CONCLUSIONS

With this background, I have arrived at the following conclusions for developing a method for AFLC to deal better with production leadtime.

In terms of a tracking device to keep AFLC informed about changes in leadtime, I would add to the present tracking system of the 12 commodities, the use of the Early Warning System. As noted in this report, this system provides a quarterly newsletter which indicates commodity groups in which

shortages are expected to occur. This information, when used with the present tracking system, should provide a management alert capability which will allow AFLC to avoid situations such as that which occurred in 1974-75.

Since this alert system only warns management of impending problems, my second recommendation is that the contractor survey system also be continued. This will provide the updating of PLT for items that have not been purchased for longer periods of time. This recommendation is based on the assumption that the system is presently being properly and uniformly implemented by all the ALCs. There have been some indications in my research that this is not in fact the case. I therefore recommend that a study be initiated to investigate ALC compliance with the contractor survey system. Part of that investigation might focus on the issue of contractors' alleged habit of overestimating PLT. Additional areas could be the thoroughness of posting done by item managers and whether different ALCs are implementing this system in identical ways.

The third recommendation would be to postpone any further research with the external regression model until the usefulness of the Early Warning System can be established. If the Early Warning System is effective, then it will provide the

same information that a regression model would give, without additional effort on the part of AFLC. In addition, the forecasting model developed was less than satisfactory, suggesting that additional effort on the part of PP or XRS would be necessary in order to establish an effective predictive device. Finally, the Early Warning System is broken down to five-digit SICs. Thus, more complete information would be provided than presently available in the 125 categories from *Purchasing Magazine*.

My final conclusion is that all external models, regardless of their accuracy will show themselves useless from an *operational* point of view. This is because of the insurmountable problem of disaggregation, or taking predictions about industry groups and transforming (disaggregating) that information down to the item level. My attempt to interface national stock classification system with the 125 product groups from *Purchasing Magazine* proved fruitless. It therefore seems obvious that an investigation of internal smoothing techniques should be undertaken next. I feel that one of the smoothing techniques described in Chapter III, or some sort of auto-regressive model might be best in predicting PLT. The data for testing such a model is available from the J014 and J042 systems. The implementation of a smoothing model into the D041 and D062 would not preclude the continuing use of the

of the contractor survey system, or the ultimate verride by the item manager who might have some current information which causes him to adjust PLT based on his own judgement.

DEPARTMENT OF THE AIR FORCE
Headquarters, Air Force Logistics Command
Wright-Patterson Air Force Base, Ohio 45433

AFLC REGULATION 84-4

28 July 1976

Production

PRODUCTION LEADTIME ACQUISITION

This regulation sets forth responsibilities and provides guidelines to the Air Logistics Centers (ALCs) for identifying current production leadtimes for use in requirements computations.

1. **General.** The Air Force Logistics Command worldwide mission requires that aerospace forces be provided logistics support, materiel, and services. Achievement of the mission requires careful planning, scheduling, and management in the development of materiel requirements. Such requirements are determined by the materiel management function. Procurement and production provides the necessary procurement support to plan for and obtain items required for mission programs. To assure effective and timely determination of requirements, materiel management requires accurate estimates of production leadtimes (PLT). PLT is continually subject to change, depending on numerous factors. Requirements computation methods generally rely on historic PLT; for example, leadtime experienced on the latest buy. On occasion, recent contractual quotes, PLT for similar items, and current contractor estimates are used. The accuracy of PLT maintained affects the determination of realistic delivery schedules and ultimately the ability of contractors to meet those schedules and fill command needs.

2. **Responsibility.** The Directorate of Procurement and Production (D/PP), Procurement Planning and Technical Support Branch (PPDM) is the office of primary responsibility (OPR) within each ALC for the acquisition of current PLT (except at Ogden ALC where the D/PP will designate the OPR). Upon request, PPDM will obtain current PLT for specific National Stock Numbered items identified by the Directorate of Materiel Management (D/MM), Requirements Branch (MMMR).

3. **Program Operation:**

a. The primary means of determining current PLT estimates is by direct written or verbal contact with the sole source or most recent contractor. Contractor response to such a request is to be voluntary and at no cost to the Government. Written requests are the preferred means of contact as they can better identify to the contractor the need for and mutual benefits to be derived from current PLT. All written requests will be made, using a letter similar in content to that of the sample letter included as an attachment hereto.

b. This regulation will be implemented through the use of the semi-automated Production Leadtime Survey. This survey utilizes an annual product generated from data in the D041, D062, J041 and J014 systems. It identifies recoverable and economic order quantity items projected to be in a buy position during the budget year that do not have a procurement action in process or completed in the last six months, or are otherwise screened by programmed edit routines. The product will normally be run in the month of March based upon the 31 December computation cycle. MMMR will forward the survey printouts, grouped in contractor sequence to PPDM for processing the survey to contractors. Computer generated contractor address labels will be forwarded directly to PPDM by AC.

c. PPDM will attach a form letter to printouts applicable to each contractor. The computer printed address labels will be used to the greatest extent possible on the form letters and envelopes for mailing the survey to contractors. In those instances where an address label is not available for a contractor, PPDM shall obtain the address. Survey printouts will be forwarded to contractors within 10 workdays after receipt from MMMR.

d. Success in accumulating realistic PLT is dependent on close cooperation with the contractor. Contacts with contractors shall emphasize that response is voluntary. Any PLT received should be in the form most convenient to the contractor. However, the survey printout is intended for use by contractors as a turnaround document for their convenience. A return envelope addressed to PPDM will be enclosed in the package sent to the contractor. Replies received by PPDM will be forwarded to MMMR within 5 workdays. If PLT for any item appears unreasonable, the IM may identify that item to PPDM for further investigation and research.

e. For those items that the IM determines the contractor quoted PLT should be entered into the requirements computation, the following time standards for file maintenance action apply: EOQ items—two weeks; recoverable items—during next quarterly computation cycle.

f. Occasionally the IM may decide that an item not included in the survey needs PLT update. In this case he should forward his request through channels to PPDM. The request should contain information substantially the same as in the survey printout (for example, NSN, supplier or MFR code, part number, item name, PLT, IM code, etc). PPDM will forward this information to the contractor as

Supersedes AFLCR 84-4, 14 Mar 74. (For summary of revised, deleted, or added material, see signature page).
OPR: PPMP
DISTRIBUTION: X

AFLC-WPAFB-AUG 76 150

an attachment to the form letter described in paragraph 3a. The contractor reply should be forwarded to MMMR in accordance with paragraph 3d—time standards set forth in paragraphs 3c, d and e are applicable to these items.

g. Should a contractor be unable to provide current PLT, or direct contact with a contractor is determined to be impracticable, a best estimate should be provided, upon IM request, based on item analysis and comparisons with leadtime for similar items, trend information available in industry trade publications, or information from other Government activities. When such estimates are given, MMMR shall be advised of the basis for same.

h. The Contract Administration and Operations Branch (PPDO or at Ogden ALC, the Contract Management Branches, PPZC and PPSC) may occasionally discover PLT changes through its production surveillance function. Significant changes should be referred to PPDM for research and forwarding of current PLT to MMMR. Updating such information may help to avoid future contract production problems by assuring more realistic schedules. Referral of such changes should, however, be restricted to unusual cases; for example, where it is known that historic PLT will not be updated for several months.

4. **Program Control.** PPDM will use a carbon copy of the survey printout as a control register to annotate progress and final disposition of each individual or collective (in the case of a list of items) request resulting from the semi-automated PLT survey. Items manually identified by IMs for update action will be recorded and controlled by a separate log.

5. **Reporting.** The Production Leadtime Survey Report has been created to measure the fluctuation of production leadtimes and to assess the value of the production leadtimes acquisition program. The report has been assigned Report Control Symbol RCS: LOG-PP(A)-7601. The report will be forwarded by mail to AFLC/PPMP with a copy to MMRRS by 30 June, with data as of 15 June. The following information will be contained in the report:

1. Number of Items Identified for the Survey
2. Number of Items Sent to Contractors
3. Number of Contractors Queried
4. Number of Items Returned by Contractors
   A. Number with Increased PLT of:
      (1) One Month
      (2) Two Months
      (3) Three Months
      (4) Four or More Months
   B. Number with Decreased PLT of:
      (1) One Month
      (2) Two Months
      (3) Three Months
      (4) Four or More Months
   C. Number Unchanged
5. Number of Contractors Responding
6. Number of Items Receiving File Maintenance Actions.
   A. Number with Increased PLT
   B. Number with Decreased PLT

HQ AIR FORCE LOGISTICS COMMAND
OFFICIAL
UNITED STATES AIR FORCE

F. M. ROGERS, *General, USAF*
Commander

C. W. MORIN, *Colonel, USAF*
Director of Administration

1 Attachment
Sample Letter

Summary of Revised, Deleted, or Added Material

The PLT Acquisition program is revised to specify the Semi-Automated PLT Survey as the nucleus of the program in lieu of manual techniques. Also, Time Standards for processing the PLT survey and an RCS rep ing requirement are added.

DISTRIBUTION: X

HQ USAF/LGPMA	2
AFISC/DPAL and AUL/ [illegible]	2 ea
HQ AFLC	4
(PPMP.....1; P[illegible]-SA.....1; IG.....1; MMRRS.....1)	
ALCs (PPDM.....5; MMMR.....5; DA.....2)	15 ea
2750/DAPL	3

SAMPLE LETTER

From: .... ALC/PPD

Subject: Request for Production Planning Information

To: (Contractor)

1. Air Force Logistics Command equipment and spare part buy determinations are based primarily on procurement experience with indivdual items. To assure effective procurement planning, it is essential to have accurate estimates of production leadtimes. Accurate estimates also benefit the supplier of items by assuring realistic delivery schedules. Inventory managers, therefore, periodically select items which require production leadtime update.

2. Our records indicate that you have previously furnished (National Stock Number, Nomenclature) to the Air Force. (Note: for a number of items, substitute: Attached is a list of items, by Federal Stock Number and Nomenclature, which our records indicate you have previously furnished the Air Force.) To assist in our planning, we would appreciate your providing current estimates of production leadtimes; for example, that time from receipt of an order to shipment of first production units, assuming an economic production run. The estimates should assume either a follow-on order, where your experience indicates periodic Government orders and continued production, or an initial order, where an item has been infrequently or irregularly purchased.

3. Current production leadtimes may be provided in whatever form you desire; however, the attached list was designed for your convenience to be annotated and returned in the envelope provided.

4. This request is strictly for planning purposes and response is entirely voluntary. It should not be considered as an indication that a procurement of the identified items is forthcoming or contemplated, or that the Government intends to pay for the information. Sole purpose for the request is to obtain accurate leadtimes which can serve our mutual interests.

5. Should you desire additional information regarding this request, please contact ....

REFERENCES

1. "Wide Spread Scarcities Haunting Industry" Aviation Week & Space Technology, Vol 100, No. 9 (4 March 1974) P. 13

2. Bernstrin, G. B. "Mean Lead Time," Applications Development Division, U. S. Navy Supply Depot, Mechanicsberg, Penn, July 1967

3. Brantley, Jim and Loreman, Donald E, "A Comparative Analysis of the D041 Single Moving Averages and other time series Analysis Forecasting Techniques," Air Force Institute of Technology, Wright-Patterson, Ohio, 1977

4. Brown, R. G., Smoothing Forecasting and Production of Discrete Time Series, Prentice Hall Inc., Englewood Cliffs N. J., 1973

5. Brown, R. G., Decision Rules for Inventory Management, Prentice Hall Inc., Englewood Cliffs N. J., 1967.

6. "Coping with Shortages," Business Week, No. 2348 (14 September 1974), pp 56-183

7. Davis, Allan S., "An Investigation of Selected Business Indications as Related to Aerospace Materials Lead Time," Air Force University, Maxwell Air Force Base, Alabama, May 1975

8. Buffa, Elwood & Taubett, William, Production Inventory Systems: Planning and Control, Richard Irwin, Homewood, Ill, 1972.

9. Green, Paul E. and Tull, Donald, Research for Marketing Decisions, Printice Hall, Englewood Cliffs, N.J., 1966.

10. Hill, Joseph Stanley, "Analysis of Production Lead Time for Missile Repair Parts Contracts Dealing with Cable Assemblies and Wiring Harnesses" USAMC Interim Training Center, Texarkana, Texas, April 1975

11. Johnston, J., Econometric Methods, McGraw-Hill Book Company, New York, N. Y., 1963

12. Kruse, Karl, "A Forecasting Order and Ship Tune for CONUS Depots," DRC Inventory Research Office, U. S. Army Logistics Management Center, Philidephia Penn, June 1977

13. Launer, Robert L., "Forecasting Contract Administration Workload", U. S. Army Procurement Research Office, Fort Lee, Virginia, June 1975

14. Morgan, J. N. and Sonquist, J.A., "The Detection of Interaction Effects," The Michigan Institute of Social Research, Monograph No. 35., 1964.

15. Meitz, A. A. and Breaux, B. C. "How to Cope with Supply Shortages," Harvard Business Review, Vol 53, No. 1 (January-February 1975) P. 91

16. Lewis, Colins P. "The Versatility of the Exponential Smoothers Portfolio of Skills," Production and Inventory Management, Vol 19. No. 2., 1978.

17. Nie, Norman et.al., Statistical Package for Social Sciences, McGraw-Hill, New York, N. Y. 1975
istration Load Time (PALT) Management and Performance Criteria," U. S. Army Procurement Research Office, Fort Lee, Virginia, March 1977

19. "Lead Time Turns Out-Landish," Purchasing Magazine, Vol 78, No. 2 (21 January 1975), P. 72

20. Requirements Procedures for Economic Ordering Quantity Items, AFLC Regulation 57-6, 26 September 1977

21. Richards, F. Russell, "On Choosing a Family of Probability Distributions for Lead Time Demand," Naval Post Graduate School Montery, Cal. February 1974

22. Short, Roger Kim, "Determining The Distribution of Procurement Lead Times using the Graphical Evaluation and Review Technique," USAMC Inter Training Center, Texarkana, TX, March 1975

23. Tersine, Richard J., Materials Management and Inventory Systems, North-Holland, New York, N. Y., 1976

24. "From Shortage of Surplus-Latest Turn for Industry," U. S. News and World Report, Vol. LXXVII, No. 19 (4 November 1974), P. 21

25. Wesolowsky, George O., Murtive Regression and Analysis of Variance, John Wiley, N. Y., 1976.

26. Wheelock, L. "Production Lead Time Forecasting," Inventory Research Office, U. S. Army Logistics Management Center, Fort Lee, Virginia, January 1972

27. Whybark, D. Clay, "A Comparison of Adoptive Forecasting Techniques," The Logistics and Transportation Review, Vol i. No. 3, 1972.

28. Yawitz, Aubrey A., "Variability of Administrative Lead Time and Production Lead Time for TROSCOM Managed High Velocity Items," Systems Analysis Office, U. S. Army Troop Support Command, St. Louis, Missouri, December 1973

29. Yawitz, Aubrey A., "Evaluation of Administrative Lead Time and Production Lead Time for TROSCOM's Secondary Items," Systems Analysis Office, U. S. Army Troop Support Command, St. Louis, Missouri, November 1973

AN EXPERIMENTAL DESIGN FOR SELECTING A FORECASTING MODEL FOR PREDICTING LRU FAILURE RATES UNDER CHANGE

BY

GORDON K. CONSTABLE

12 JUNE 1978 - 18 AUGUST 1978

Prepared For

USAF-ASEE Summer Research Faculty Program
Air Force Office of Scientific Research
Washington, D. C.

And

Directorate of Management Sciences (XRS)
Deputy Chief of Staff/Plans and Programs
Headquarters Air Force Logistics Command
Wright-Patterson Air Force Base, Ohio 45433

Gordon K. Constable
Wright State Univ.

ABSTRACTS

This study involves the design of two experiments for predicting failure rates for line replaceable units (LRUs). The first chapter outlines the objective of the study and points out the problem with the model currently in use. Chapter II is devoted to discussing the models currently proposed in the literature and to the variables felt to affect the LRU failure rate. The third chapter presents the "ideal" experimental design utilizing the variables discussed in Chapter II.

Chapter IV considers the experimental design for determining a model for predicting LRU failure rates based on utilizing the data available. The final chapter summarizes the report and makes a recommendation on implementing the experiments.

PREFACE

This study was undertaken as part of the USAF-ASEE program during the period 12 June 1978 to 18 August 1978. The study objective was to design two experiments for determining a model for predicting LRU failure rates. The first design assumes that data for any variable thought to affect LRU failure rates will be available. The second design is concerned primarily with determining a model for accurately predicting the number of LRU failures without regard for the underlying causes. However, the model must be able to predict well for changing senarios for operating aircraft.

The report reviews several models proposed for predicting failure rates and discusses variables given by reports and coming from discussions with individuals involved in flying and maintenance operations. The designs of the two experiments are presented and several suggestions made for implementing the experiments.

The author of this report extends thanks to Mr. V. Presutti (XRS) and Mr. Maurice Shurman (Boeing Company) for their assistance during the research and preparation of this report. In addition, I would like to thank Mr. Don Casey, Mr. H. D. Hunsaker, and most of the other personnel in XRS for freely giving of their time and helpful advice.

TABLE OF CONTENTS

EXECUTIVE SUMMARY

1. In response to requests from Headquarters AFLC, this study was undertaken by a professor participating in the USAF-ASEE Summer Faculty Research Program in 1978. The work on this project commenced 12 June 1978 and continued through the 18 August 1978. The central question addressed is whether a more accurate method of forecasting line replaceable unit (LRU) requirements might be employed for predicting War Reserve Material (WRM) requirements than the current method. The current model assumes that LRU failures are directly proportional to the total number of flying hours. This implies that the number of spares required for twelve sorties of four hours each would be the same as required for four sorties of twelve hours each. A number of studies have been done which indicate that the number of failures is a decreasing function of the sortie length. The main concern of this study is to design an experiment which will test a number of models for determining the LRU requirements when the policy parameters change, i.e., number of sorties and the lengths of sorties. As part of the objective of the study, the experimental design should be capable of generating a model for predicting LRU failures that could possibly be used by the Item Managers in determining their requirements.

2. The current models which were reviewed include flying hours, sortie, flying hours and sortie, time dependent, and a number of time series. The flying hours and sortie models have both been found to explain about .60 proportion of the variance in the number of maintenance actions (Casey [7]). When sorties and flying hours are combined, the explaination of the variance rises to .66 (Casey [7]. The time dependent model appears to do much better with percentage errors quoted from one to fifteen percent. However, none of these models have been applied to predicting LRU failure rates. The time dependent model has only been applied to a few specific cases using aggregate data, but the results have been impressive (Shurman [32]). In addition to these models which have been proposed for determining the number of failures or maintenance actions, a number of reports have indicated that a variety of other variables affect failure rates. The most important of these variables appear to be the type of aircraft, the type of system, the mission the aircraft flies, and the utilization of the unit. A number of other factors have some affect on the LRU failure rate such as: the aircraft crew, the maintenance personnel, and the attitude of these two groups towards each other. However, since these variables are difficult to quantify and they are dynamic, they are not considered as part of the analysis of this report.

3. In developing the ideal experimental design, the variables indicated as affecting the LRU failure rate were included in the analysis. The assumption of the design was that the data for each variable could be obtained. Given the number of variables and the number of potential values for each variable, the design of the experiment includes the reduction of the initial set of variables. Regression is specified as the appropriate tool for specifying the relationship between the LRU failure rate and the reduced set of variables. Non-linear and interaction relationships are considered by the experimental design.

4. A second experimental design has been developed as a result of the potential problems that might be encountered in implementing the ideal experimental design. One of the primary problems is the amount of time required before the results can be obtained. The reason for the time requirement is the need to collect new types of data in large quantities. Thus, the modified experimental design utilizes existing data for the analysis. The proposed models using flying hours, sorties, and time dependent data can be statistically tested and compared to see which model performs better than the others. The procedure for collecting the data is outlined, the analysis the data should be subjected to is given, and the tests for comparing the models are discussed.

5. The advantage of employing the modified experimental design is that if a better model is determined, immediate improvement

can be made in predicting LRU requirements particularly WRM requirements. This information could also be entered into the D041 system. If the ideal experimental design is completed, the underlying causes of the LRU failures could be determined. The results might possibly affect decisions in scheduling aircraft for missions, preventive maintenance actions, scheduled overhauls, inventory policies among others. Here the results could be far reaching with the long term benefits very significant.

CHAPTER I
INTRODUCTION

1. General

The objective of this study is to develop an experimental design for evaluating a number of models for predicting spares requirements for line replacement units (LRUs). The models will predict the "failure rate" of LRUs as opposed to the actual spares required. Once the failure rate is determined--in reality actually a demand rate--it can be put into the current stocking models to determine the number required by the system. As part of the objective of this study, the model for predicting the failure rate of an LRU must be able to do so under changing conditions; i.e., a change in sortie length.

In order to clarify the difference between the objective of this study and the current literature, I will define:

Failure Rate - Refers to the malfunction of a unit on an aircraft. The malfunction may require an adjustment; a remove, fix, and put back; a remove and replace with the unit being thrown away, condemned, or repaired. This definition corresponds to the use of failure rate in most of the literature.

LRU Failure Rate - Refers to the malfunction of an LRU which results in its removal from the aircraft and replacement by a similar LRU. This definition requires an LRU failure to result in a demand on the system.

For a number of years, the United States Air Force (USAF) has been using the assumption that the failure rate is exponentially distributed. This implies that the failure rate remains constant through time. Thus, the same number of failures would be expected to occur for 12 flights of four hours each as would

occur for four flights of 12 hours each (each giving a total of 48 flying hours). While most of the people involved would readily agree that this assumption is not valid, it has provided a straightforward, inexpensive model for predicting spares requirements.

During the past ten years, empirical data have surfaced in a number of reports indicating a relationship between the failure rate and the time into a specific flight. Generally, the failure rate is relatively high for the initial stages of a flight, levels off at some relatively low rate during the cruise portion of the flight, and increases during the final stages of the flight. Figure 1.1 has graphs of failure rates versus time-in-flight taken from the listed references.

There are a number of intuitive reasons as to why this might be the appropriate pattern for failures.

1. The initial power surge into the components may overload it and cause a failure to occur. As the system continues to operate, it reaches standard operating conditions and the failure rate would stabilize.

2. The environmental control systems, such as air conditioning and heaters, may not have stabilized before the other systems have been switched on.

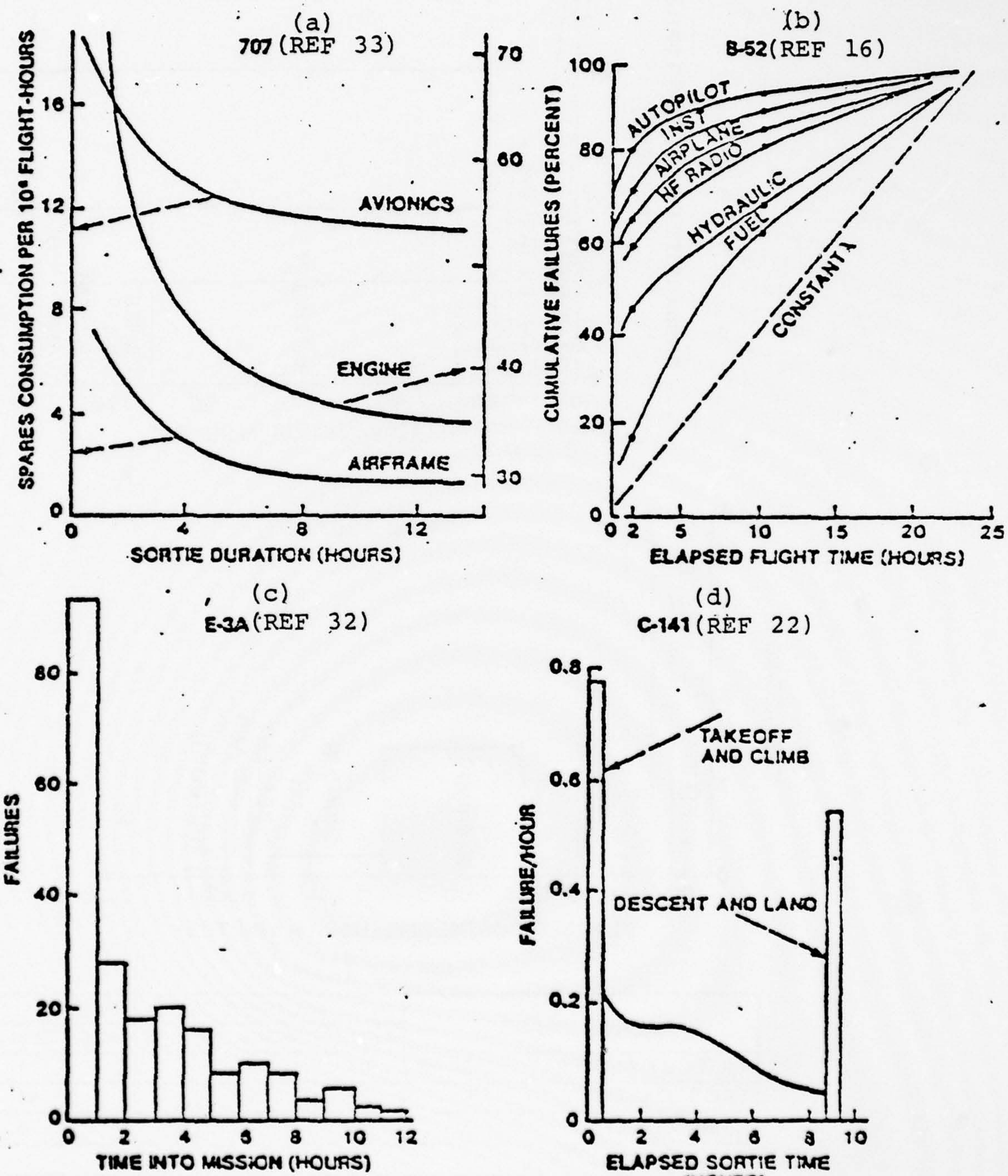
3. Low frequency vibration and shock during the taxi and takeoff increases the probability of a failure. This would be particularly true if the aircraft is loaded to capacity.

4. During the cruise (in and out) phase at a flight, most of the units would be subject only to wear-out failures since generally no special stresses exist for the unit.

5. During the landing phase, the failure rate would increase from the cruise portion of the flight because of the sudden shock from the touchdown and the increased

FIGURE 1.1

FIELD EXPERIENCE DATA*



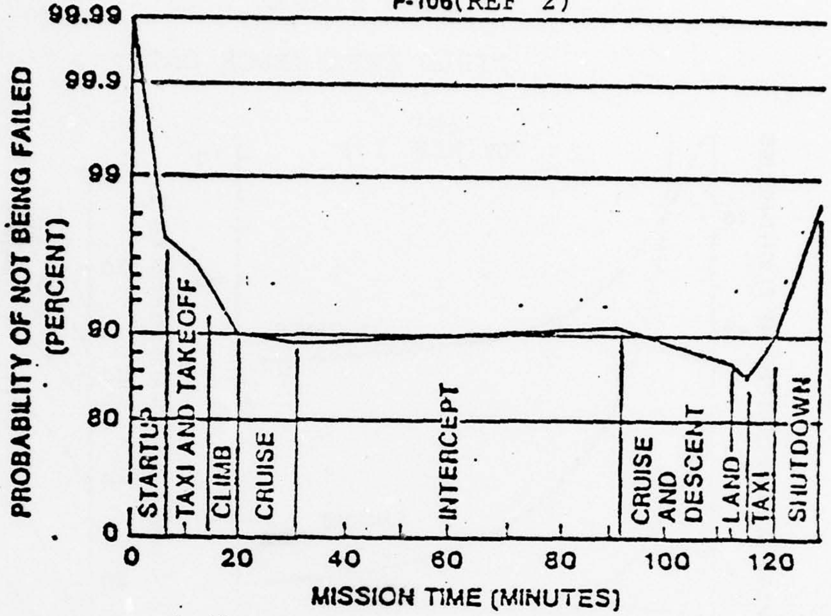
* These graphs were taken from Shurman (32, pages 3 and 3A) who copied them from the original references listed on the graphs except for part (c) which was an internal Boeing communication.

FIGURE 1.1

.FIELD EXPERIENCE DATA (CONT.)

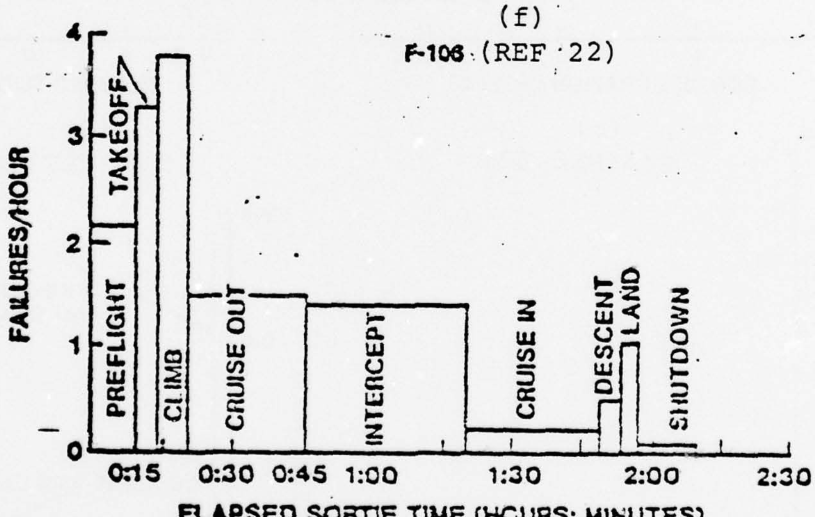
(e)

F-106(REF 2)



(f)

F-106 (REF 22)



(g)

F-106 UHF (REF 3)

SORTIE PHASE	NUMBER OF FAILURES			
	ABORT	DEGRADE	NUISANCE	TOTAL
TAXI OUT AND FIRST QUARTER	12	37	22	71
SECOND QUARTER	11	19	11	41
THIRD QUARTER	2	5	2	9
FOURTH QUARTER	3	5	3	11

vibration from keeping the engines ready in case the landing must be aborted.

In addition to the relationship between time into a sortie and the failure rate, a number of other variables have exhibited an effect on the failure rate; i.e., type of aircraft, type of system. As a natural result of these studies, a relevant question becomes: Which of these variables should be included in the failure rate model?

2. Purpose and Scope. The objective of this study is to design an experiment which will determine which variables "significantly" affect the LRU failure rate. In addition, the design should permit a variety of models to be tested to determine the "appropriate" model to be used.

Note 1. A variable may be significant but still not explain much of the variation in the data.

Note 2. An appropriate model may not be the most accurate model but may be one which is relatively accurate and requires relatively little information. (Provides a good estimate of relatively low cost.)

3. Outline of Study. In Chapter II, there will be a discussion of several reports and the models they present for predicting failures. In addition, the variables thought to affect the failure rate will be listed. The experimental design for analyzing all of the variables (assuming data available) will be presented in Chapter III. The modified design based on the data available and limited to specific models will be discussed in Chapter IV. In Chapter V, a summary of the report and a recommendation on a specific course of action will be presented.

CHAPTER II
MODELS AND VARIABLES

1. Models

There are a number of reports concerning several models for predicting failure rates. Many of these models do not predict actual failures but use as a proxy variable unscheduled maintenance man-hours, number of maintenance actions, or number of maintenance write-ups. A proxy variable is often used because of the lack of data for the desired variable and the availability of data concerning the proxy variable.

Flying Hours Model:

As previously mentioned, the flying hours model assumes that the number of spares required is directly proportional to the total flying hours. If total flying hours double, then the total spares requirement doubles. Under peacetime conditions, this is not a bad assumption. In general, any change is proportional to what was done before. The mix of sortie lengths and mission types remains relatively constant, therefore you would expect the spares requirement to be proportional. However, consider a situation (such as wartime) where the number of total flying hours quintuple, the number of sorties doubles, and the mission mix is radically altered. Here, it is not so clear what happens to the spares requirement. Casey [7] ran a number of regressions for determining maintenance write-ups as a function of flying hours for the C-5 aircraft using two digit work unit codes (WUC). The coefficient of determination (R^2) was less than 0.60 in every case. Donaldson

and Sweetland [11] reviewed data concerning the B-52, F-100, F-4C, F-5A, and the C-130 and concluded that unscheduled maintenance man-hours (UMMH) were not proportional to flying hours. However, they did find that a significant relationship existed between UMMH and flying hours even though it was not directly proportional.

Sortie Model:

The sortie model generally assumes that the increase in the number of sorties would require a proportional increase in maintenance man-hours or maintenance actions. If the additional sorties are reproductions of what has been occurring, there will likely be a proportion increase. However, if the number of sorties are held constant while the total flying hours are doubled, the number of maintenance actions or maintenance man-hours would be expected to increase. Casey [7] found that sorties explained less than .62 of the variation. In all but two of the cases he investigated, sorties gave a higher R^2 than did flying hours. By contrast, however, Hunsaker, et al, [17] found that for the F-4 fighter aircraft, it appeared that the number of maintenance actions required went up for an increase in the number of sorties when the total flying hours were held relatively constant. They were unable to test the hypothesis of a proportional increase statistically, however.

Flying Hours and Sortie Model:

Since we have two variables which explain 50-60 percent of the variation in maintenance actions required, it would appear that by combining them, we would explain most of the variation.

The drawback in this approach is that flying hours and number of sorties are directly related to each other by the average sortie length. Knowing any two of the three (flying hours, number of sorties, and average sortie length), we can calculate the third. This implies that the amount of variation in maintenance actions explained will probably not increase much by using flying hours and number of sorties together in the same model. Casey [7] found this result for the C-5 data. The largest increase in the R^2 obtained by adding a second variable (either sortie or flying hours) was 0.070. The highest R^2 was for unscheduled maintenance write-ups (excluding WUC 01-09). In this case, the $R^2 = .667$, which is not as high as one would like for prediction purposes. No F values for the significance of the regressions were given. Dow and Schnee [13] did a similar study involving B-52H engine failures. They achieved much better results than Casey. However, they added a large number of variables to their regression equation making their results subject to question. A validation on a subsequent set of data would have been helpful.

Time Dependent Model:

A number of reports have surfaced in the past ten years containing data which indicates that the failure rate is a function of time into the sortie. References [6], [7], [22], [32], and [33] indicate that the failure rate decreases monotonically as time into a flight increases. Graphs of the failure rate in a number of these references were previously shown in Figure 1.1.

Martinez [22] divided the sortie into phases: Before Flight, Takeoff, Climb, Cruise (outbound), Intercept, Cruise (inbound), Descent, Land, and Shutdown. Figures 1.1d and 1.1f demonstrate the time dependency found in this reference. The failures could only be tied to flight phase and not by time into the sortie. Thus, the probability of a failure was taken as constant for each phase. The simplicity of this model makes it seem attractive. However, for determining the aggregate spares requirement, the times for each phase of a sortie would have to be summed and then multiplied times the phase failure rate. This would be too cumbersome a procedure to use--particularly for LRUs.

Shurman [32] recommends a time dependent failure function which gives the failure rate as a continually decreasing function of time into a sortie. He develops the form of the failure rate function based on the curves presented in Figure 1.1. The failure rate function is:

$$\lambda(t) = \frac{0.45\,\lambda o\ T}{t+0.08}$$

where: $\lambda(t)$ = probability of a failure of time

λo = steady-state in-flight failure rate

T = nominal unrefueled sortie length for which the aircraft was designed (in hours)

t = time into sortie (in hours)

One advantage of this equation is that by using integration the mean number of failures for particular length sorties can be determined. Given a particular scenario specifying the

number of sorties of each length, it becomes a simple matter to determine the spares requirement. This assumes that the appropriate constants have been determined and that the model provides a good fit to the failure data. Shurman does apply the model to some additional data, and the model appears to provide an adequate fit [32, pages 14-24].

Other Models:

There are a number of other models used to predict failure rates (or more appropriately demand rates) for LRUs. Models using exponential smoothing, moving averages, K factors, and Kalman filters are based primarily on projecting the past performance into the future. Hence, the future is like the past. This is contrary to the objective of this experimental design where we want to be able to predict the LRU failure rate for a change in policy parameters; therefore, a causal model seems to be in order.

2. Potential Variables. A number of potential variables has been specified without regard to the existence of data concerning the variables. These variables have been selected based on relationships found in reports and from discussions with individuals involved. For a complete analysis of the variables, it may be desirable to collect the necessary data. The variables of primary interest are those associated with the LRUs since our objective is to predict the demand rate (LRU failure rate) for LRUs. It may be desirable to collect data associated with systems and the aircraft for consideration with more aggregate models. The variables can be classified in general

as those associated with the LRU, with the system (or subsystem), with the aircraft and with the particular sortie flown.

LRU Variables:

1. Grade of the LRU: There have been some indications that units supplied by different manufacturers do not necessarily exhibit identical operating characteristics. In some cases, units supplied at different times by the same manufacturer have exhibited different operating characteristics often by design. It might be desirable to equate grade with the number of times the unit has been repaired.

2. Complexity of the LRU: This is a measure of the unit's susceptibility to a failure occurring. It could be a measure of the number of parts in the LRU, the number of Shop Replaceable Units (SRU) or the number of ways an LRU can fail.

3. Number of cycles on the LRU: Some units wear more from being switched on and off than they do from operating over extended periods. The count probably should be started over when the LRU is repaired.

4. Operating hours on the LRU: This measure would give an indication at the operating life and its expected value would approach the MTBF if the time is initialized to zero for each repair cycle.

System (Subsystem) Variables:

1. Complexity of the system: Aircraft systems vary widely in the number of component parts and in the types of ways the system can fail.

2. Number of cycles on the system: This relationship is not as direct as it is for an LRU, but as an indication of age, it may be significant. It probably should be a lifetime counter since a large number of the component parts would have been individually rapaired at different points. If there is a general overhaul of the system, the counter could be set to zero at that point.

3. Operating hours on the system: This would measure the operating life of the system and should somewhat relate to the wear out rate. Again, since LRUs could have a variety of initial start points in any one system, this is not as direct a relationship as it is for an LRU.

4. Type of system: As shown by data in several of the references, different systems exhibit different failure rates.

Aircraft Variables:

1. Type of aircraft. The data indicate that failure rates are different for various types of aircraft. This could also be a function of the type of mission each type of aircraft generally flies.

2. Complexity of the aircraft: Some aircraft have extensive avionics, etc., and this would have an impact on the failure rate.

3. Number of sorties flown by mission type: This indicates the age of the aircraft by times used. Doing it by mission type permits different stress requirements to have an impact.

4. Operating hours by mission type: This indicates the amount of time the aircraft was exposed to the stress of particular mission types.

5. Utilization of the aircraft: A number of situations have been observed where the failure rate is affected by the amount of time between sorties. The longer the time between sorties, the higher the number of failures.

Sortie Variables:

1. Mission for this sortie: The mission would indicate the stress on the aircraft. It may be desirable to distinguish between missions of the same type, e.g., one with high altitude flight and the other a low altitude flight.

2. This would give the expected number of exposures for the LRU, the system or the aircraft to failures.

3. Operating hours expected on the unit this sortie: This would indicate the time the LRU, the system or the aircraft would be exposed to a failure.

4. Capacity utilization: Mission for which the aircraft is loaded to capacity put higher levels of stress on the unit for longer periods of time.

5. Length of sortie in hours. The length of the sortie indicates the amount of time the unit is exposed to an in-flight failure.

6. Time into sortie that failure occurred: If the failure rate is time dependent, these data will be required to estimate the time dependent failure rate function.

In reviewing the reports, there were a number of situations described which indicated that other variables may affect the the failure rates. Aircraft flying identical missions out of different bases exhibited different failure rates for identical systems. These differences have been attributed to the environment, to the crews, to the maintenance personnel, to the attitude of the crew and maintenance personnel towards each other, to the understanding of the maintenance policies by the crew and maintenance personnel, and to the individual aircraft itself.

These potential variables were omitted from the list because of the difficulty in defining and quantifying them. These differences can be dealt with by confounding them; i.e., use worldwide missions to confound weather and base effects, or by stratifying the data and holding the particular variable constant.

The next chapter will discuss the data requirements for these variables and the steps to follow in analyzing the data requirements.

CHAPTER III

IDEAL EXPERIMENTAL DESIGN

In the last chapter the variables which were thought to affect the LRU failure rate were listed and briefly discussed. In this chapter the form of the data, the amount of the data required, and the design for analyzing the data will be discussed.

1. Data Requirements

Most of the variables specified previously may be expressed quantitatively with the possible exception of the "grade of the LRU" and the "type of mission this sortie." Since many of the variables can take on a large number of numerical values, it would simplify the data collection and the data analysis in a number of cases if class intervals were set up. For example: consider setting the complexity of the LRU by counting the number of shop replaceable units (SRUs) contained in the LRU. We can set the categories of simple, normal, and complex to refer to LRUs with 0-10, 11-20, and 20 or more SRUs respectively. These categories could be considered levels or treatments in the analysis or they could take on numerical values based on the midpoints of their respective intervals. If the actual number for the average number of SRUs is determined for each LRU, the average number

of SRUs contained by all of of the LRUs in the interval could be used as the value for the category. For some of the analyses we will consider, the actual numerical values would be preferable. An example of a typical data sheet using the variables listed in Chapter II is shown in Table 3.1.
By using the variables listed in Chapter II, each sortie in essence becomes an individual data point with the dependent variable indicating that on this sortie an LRU failure occurred or did not occur. To translate the data into a failure rate requires that a number of points with identical values on the variables be collected and the number that failed be divided by the total number of observations.
This results in two complicating factors. First, the number of observations required to be reasonably confident of the accuracy of the estimate of the failure rate is large. Secondly, the large number of possible combinations of variables and the inability to control the values of some of the variables makes the possibility of exact replications unlikely.

Sample Size:

Since the proportion of failures (an estimate of the failure rate) is a binomial type of process, a standard formula can be used to calculate the sample size required for a specified

TABLE 3.1

TYPICAL DATA SHEET

Observation Number	LRU fails on flight (1) No failure (0)	Complexity of LRU	Grade of LRU	Cycles on LRU since last repair	Operating hours on LRU since last repair	Expected cycles on LRU this sortie	Expected operating hours on LRU this sortie	Complexity of System	Cycles on system since last overhaul	Operating hours on system since last overhaul	Expected cycles on system this sortie	Expected operating hours on system this sortie	Complexity of aircraft	Number of sorties by mission type			Operating hours by mission type			Length(hrs) of this sortie	Mission for this sortie	Proportion of capacity load at takeoff	Utilization of aircraft	Time into sortie failure occurred	Proportion of time into sortie failure occurred
														R	T	C	R	T	C						
1	0	10	A	54	46	2	2	25	150	50	3	4	100	10	25	100	50	25	800	4	R	.60	.2	-	-
2	0	10	A	61	52	3	4	25	162	60	5	3	100	8	35	120	45	140	720	5	T	.70	.4	-	-
3	1	10	B	22	30	0	0	25	134	25	2	3	100	4	16	64	20	80	640	10	C	.85	.3	.8	8/10
4	0	10	C	40	45	2	2	25	220	70	2	2	100	10	15	60	20	30	120	6	T	.75	.6	-	-
5	0	10	B	30	20	4	4	25	107	60	4	6	100	-	-	150	-	-	300	4	T	.60	1.2	-	-
6	0	10	C	50	37	3	3	25	54	36	2	5	100	2	10	24	8	60	200	8	C	.90	1.0	-	-
7	1	10	A	25	14	6	4	25	190	120	3	7	100	5	45	50	25	360	600	10	C	.90	.7	.2	2/10
8	0	10	B	40	38	2	3	25	150	107	4	4	100	10	30	180	30	120	1080	20	C	.95	.5	-	-
9	0	10	C	56	50	3	2	25	75	28	3	2	100	2	12	16	5	36	64	15	C	.78	.4	-	-
10	0	10	C	17	10	4	5	25	64	90	4	4	100	33	24	12	10	48	48	3	R	.55	.7	-	-
11	0	10	B	25	22	1	6	25	100	80	2	7	100	2	60	80	10	120	1500	24	C	1.00	.8	-	-

accuracy at some confidence level. The formula for calculating the sample size is shown below:

$$N = Z^2 P(1-P) SP^2$$

Where: N is the number of observations required.

P is the true proportion of failures.

S is the $\pm$ interval about the value of P.

Z is the number of standard deviations required for a specific confidence level, i.e., Z equal to 1.96 for a .95 confidence level.

For example:

To estimate p for an LRU that is suspected of having a failure rate of 0.001 and an accuracy from .0005 to .0015 at the .95 confidence level, a sample n equal of 15,351 observations is required.

$$N = (1.96)^2 \ (0.001) \ (1-0.001)/(0.005)^2 = 15,351$$

By way of contrast for a relatively high failure rate of .2 at the same relative accuracy (½ of p), the sample size required becomes 61.47 or 62 observations for the same confidence level.

$$N = (1.96)^2 \ (.2) \ (1-.2)/(.1)^2 = 61.47$$

AD-A065 651 OHIO STATE UNIV RESEARCH FOUNDATION COLUMBUS F/G 1/3
USAF-ASEE (1978) SUMMER FACULTY RESEARCH PROGRAM (WPAFB). VOLUM--ETC(U)
NOV 78 C D BAILEY F44620-76-C-0052
UNCLASSIFIED AFOSR-TR-79-0232 NL

5 OF 5

AD
A065651

END
DATE
FILMED
5-79
DDC

This indicates two actions that can be taken to reduce the sample size requirements. First, the analysis can concentrate on LRUs with relatively high failure rates, or secondly, the unit considered can be aggregated to a larger unit (system or subsystem) which will result in a higher failure rate. However, in the latter case the specific LRU failure rate can no longer be determined directly but must be infered as a proportion of the system or subsystem failure rate.

2. Initial Analysis

Initially we would like to test which variables significantly affect the failure rate without inferring the specific form of the relationship. Analysis of variance (ANOVA) is particularly suited for this task. A complete explanation of ANOVA is too detailed to include here, particularly when we consider the data requirements for a complete balanced ANOVA design. For additional information see Neter [28,419+] and Ostle [29,279+].

In ANOVA the variables would essentially be treatments with a different values on each variable considered as a level of the treatment. A continuous variable would require

a treatment level of each unique value in the data. In general a large number of levels would result in too many unique combinations of the variables, therefore the number of levels are reduced by setting up class intervals containing a range of values. For determining the LRU failure rate, from initial consideration with the possible exception of the type of aircraft which will be held constant for the sample data. There are four variables associated with the LRU and six variables associated with the sortie for a total of ten variables. If each variable is restricted to three levels or class intervals, there will be 310 or 59,049 unique combinations of levels for the ten variables. Since the values for some of the variables cannot be controlled, obtaining specific combinations may be difficult if not impossible. One should note that if a sample of 59,049 unique data points, is obtained, there would not be any combinations having a repeat sample or replication. This is not as bad as it first appears since many of the higher level interaction terms could be used to estimate the error term. An initial advantage of employing the ANOVA would be the ability to test for interaction and non-linear relationships if the suspected non-linear relationships could be specified prior to analyzing the data. In conclusion, the ANOVA is not a feasible technique for use at this point in the analysis.

There are three techniques that can be used to select a reduced set of variables for further consideration. They are factor analysis, discriminant analysis and automatic interaction detection (AID). Discriminant analysis assumes a linear model for predicting group membership while factor analysis and AID do not make a specific assumption of an explicit model. In fact, factor analysis does not utilize the independent variable (LRU failures in our case) in analyzing the data. These techniques will be briefly discussed with references listed for examples and a more detailed explanation of each technique.

Factor Analysis [14], [15], [30]:

Factor analysis takes an n by m matrix (N-observations, M-variables) and searches the data matrix looking for patterns for the m variables. It redefines the matrix by essentially combining or grouping similar variables. The factors are chosen in such a way as to insure that the first factor explains as much of the variation in the data as possible. The second factor is chosen because it explains the maximum amount of the remaining variation in the data. The process is repeated until a cutoff point is reached i.e., 95 percent of the variation explained, or until a specified number of factors has been defined. Factor

analysis is most useful when some of the m variables are similar or highly correlated. The output of factor analysis is a factor matrix which lists the factors and the loadings of the m variables on each of the factors. Table 3.2 contains a n example of a factor matrix. If four of factors contain 80-90% of the variation in the matrix, four variables, each loading high on one of these factors, can be used as substitutes for the factors. The result is a reduced set of variables which are relatively independent of each other and still capture a majority of the variation in the data matrix. In general, these four variables would more likely to explain or predict the value of another variable than any at the remaining variables for this data set. For example, the Table in 3.2 variables 1, 2, 4 and 10 load on factor 1, variables 3 and 7 on Factor 2, variables 8 and 9 on Factor 3, and variables 5 and 6 on Factor 4. For additional analysis of the data, variable 1 can be used as a substitute for Factor 1, variable 7 for Factor 2, variable 8 for Factor 3 and variable 6 for Factor 4. This provides a reduced set of variables which can be used for further analysis. For example, a number of unique combinations for four variables each having three possible values is 34 or 81 combinations--considerably less than the 59,049 combinations for ten variables having three value each.

TABLE 3.2

FACTOR LOADINGS

(HYPOTHETICAL)

Variable	1	2	3	4
1. Grade	0.98	0.01	0.01	0.00
2. Complexity	0.76	0.10	0.02	0.05
3. Number of Cycles	0.03	0.67	0.04	0.16
4. Operating Hours	0.55	0.30	0.01	0.06
5. Expected No. of Cycles	0.05	0.03	0.10	0.68
6. Expected Operating Hours	0.02	0.01	0.03	0.85
7. Aircraft Utilization	0.01	0.87	0.02	0.05
8. Proportion of Capacity	0.02	0.03	0.91	0.01
9. Sortie Length	0.01	0.05	0.88	0.05
10. Mission	0.76	0.20	0.01	0.02

Discriminant Analysis [14], [26], [27]:

Discriminant analysis is a predictive technique which assigns an item to a group based on the item's values on the number of characteristics (variables). The prediction is based on an index value determined by substituting the item's values on the variables into a functional relationship (equation) and comparing that to each group's average index value. In the case of LRU failures there are two natural groups-sorties on which no LRU failures are reported and sorties on which an LRU failure is recorded. The advantage of employing discriminant analysis is that it will indicate which variables account for the largest differences between these two groups, i.e., the variables that most effect the occurence of an LRU failure.

In addition, discriminant analysis provided an ability to examine the data in a variety of ways. Since there is not much known about the form of the data, the ability to examine the data in different ways could provide additional, useful information. The functional relationship that results from the analysis could be used to predict rather a sortie which is planned belongs to the failure or non-failure group for a particular LRU. This would be useful in determining when to do preventive maintenance on the LRU, or

in planning for critical missions where a high probability of success is desirable.

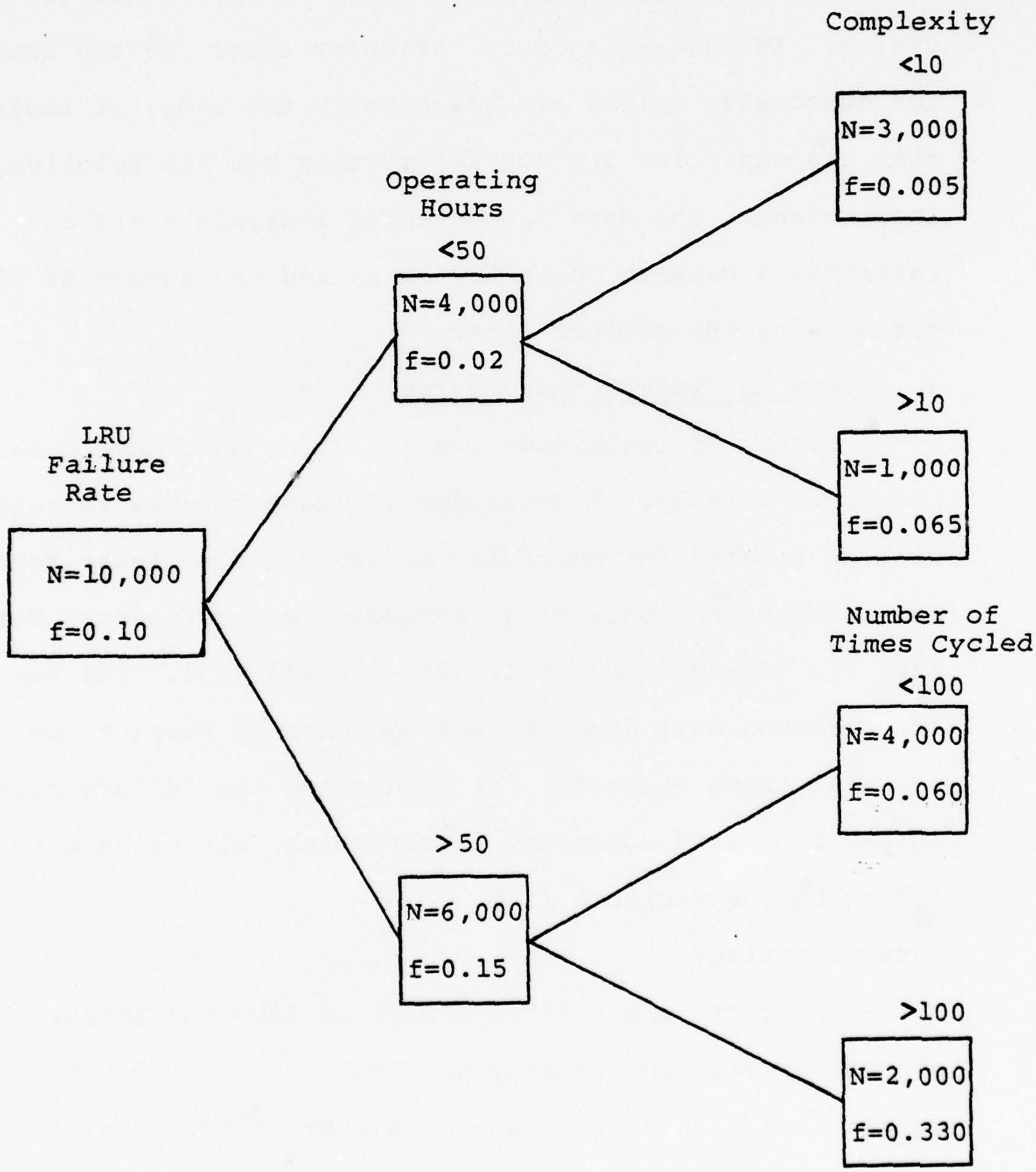
Automatic Interaction Detection (AID) [14], [24], [25]: AID is a technique for determining complex, interactive relationships in data. The general principal of the program is the application of a prestated strategy searching for those predictors that increase the ability to account for the variance in a dependent variable. The AID technique uses a repeated one-way analysis of variance techniques to explain as much of the variation of the dependent variable as possible. The search procedure performs a sequence of two-way splits of the data sample, choosing those splits (and the variables) that separate the sample into two subgroups that maximally will account for the remaining variation in the dependent variable at each stage of the analysis. At each of these stages it will consider all allowable splits on each predictor (independent) variable. While AID can identify potential relationships, it is questionable on the matter of statistical significance and further analysis should be done.

The analysis done by AID can be depicted graphically by a tree diagram as shown in Figure 3.1. In this hyphothetical case, we can see the interaction of the number of hours an

FIGURE 3.1

AID EXAMPLE

(HYPOTHETICAL)



LRU has operated and the number of times it has been cycled (turned on and off). In the case of low operating hours (less than or equal to 50), the complexity of the LRU becomes a more important indicator than the number of cycles. If the sequence of variables along the two branches for successive splits are essentially the same, it implies that the variables are not interacting but are relatively independent. The data as presented indicate a strong interaction between operating hours and the number of times cycled with the failure rate.

3. Steps in Experimental Design

A number of techniques for analyzing data have been discussed briefly. A knowledge of least squares regression will be assumed for multiple, binary and non-linear regression. For information concerning regression see references Neter [28, PP. 178-407] and Ostle [29, PP. 159-205]. The rest of this section will describe the sequence of steps to be taken in determining the model for predicting the failure rate for an LRU from data collected on variables felt to have an effect on the variable rate.

Data Selection:

Since there are a large number of LRUs the initial experiment will be for only one LRU. If the analysis proves useful then it can be extended to other LRUs. The LRU selected should have a number of characteristics. First, it

should be one where there is significant interest in the results. This implies that it must have a relatively high cost and a relatively high LRU failure rate. It should be a unique part which implies that it is only used on one type of aircraft and that the number of bases having this aircraft should be limited to facilitate the collection of data. There should only be one of this LRU on each aircraft. It would be beneficial if most of the variables listed have a limited number of possible values, but that these values be distinctly different: for example, two or three mission types and sorties of four eight and twelve hours. This would limit the number of variable combinations and make the class intervals for continuous variables easier to set.

The number of observations to collect can be determined by using the formula discussed in section 1 of this chapter. However, the LRU failure rate has to be based on a sortie, number of flying hours, time into the sortie, or some other parameter. The unit chosen should relate to the models that will be used for the prediction, i.e., LRU failure rate per sortie. In any event the data concerning several thousand sorties should be gathered.

Data Collection:

The collection of the data must be done with care to ensure consistency and accuracy. If a number of bases

are involved, it may be beneficial to rotate crews and aircraft to neutralize their effect on the LRU failure rate. If only one base is involved, it would provide a more controllable situation but would lengthen the time required to collect the required quantity of data.

Since the model is being derived from the data, it should be validated. One of the best procedures for doing validating the model would be to gather enough data to have two samples. Once the model is derived, it can then be tested on the remaining sample. Since the model developed is to be used under changing conditions, the second sample should not have the same average values for the variables as the sample used for deriving the model. It would be helpful to have a significant difference between the two samples on variables such as the number of sorties flown and the average sortie length. It would be possible to collect the data for the second sample while the data for developing the model are being analyzed.

Initial Investigation:

Factor analysis should be used in the initial screening of the data in order to reduce the number of variables under consideration. However, since little is known about the data being collected, it may be advisable to run both discriminate analysis and AID. The binary nature of the original data fits the discriminate model perfectly. This

analysis can tell you which variables are important in making the distinction between failures and non-failures. Since factor analysis does not relate the variables to the LRU failure rate, the proxy variables for each factor should be chosen based more on the variables relationship to the LRU failure rate than to its loading on a particular factor.

AID can indicate that the interaction between two variables is important in determining the failure rate. This may modify the form of the model that should be used on the data, i.e., AID may indicate that an interaction term for two variables should be added.

Variable Selection:

If only factor analysis is used the variables should be selected based on the highest loading for each of the initial factors explaining a majority of the variation in the data. If discriminant analysis used, the variables selected as proxies for the factors should be those that discriminate between failures and non-failures even though they do not load the highest on a particular factor.

Once the reduced set of variables has been selected, the reduced set of data should be sorted to determine if each unique case represented has enough replications to provide relatively accurate LRU failure rate. If there are not enough replications, then the measurements on the reduced

set of variables will have to be grouped into class intervals to provide more replications. If necessary, more data can be gathered recording the reduced set of variable values only.

Data Transformation:

The Data must be transformed into LRU failure rate data. This can be done by collecting all sorties having identical values on the reduced set of variables. The number of LRU failures that occured must then be found and divided by the total number of identical sorties to determine the failure rate. This will significantly reduce the number of data points. If there are four variables with three levels each in the reduced set, there would be at most 81 unique combinations of variables if all combinations are present in the data.

Model Development:

Multiple linear regression will be used on the LRU failure rate data for determining the prediction model. If AID was used in the initial analysis of the data, it may be desirable to include an interaction term if one is so indicated. The regression can be run with an without the and without an interaction term to see if incremental gain for the interaction term is worth the effort.

The results of the regression should be tested for statistical significance. This implies using a T-test on

the coefficient terms to see if they are significantly different than zero, and using the F-test to see if the total regression is significant. Then a subjective judgment must be made concerning the coefficient of determination (R^2) associated with the regression. Is the R^2 satisfactory or should the model be modified in an attempt to improve the R^2? If the decision is made to improve it, non-linear components can be added to see what improvement results. Non-linear relationships for sortie length or time to failure might prove most beneficial since these relationships have been shown to be non-linear. The process of testing additional relationships can be continued until an adequate R^2 is obtained or it is felt that these variables will not provide an appropriate LRU failure rate model.

Note, however, the model needs to be checked beyond the R^2 value for predicting each combination of the significant variables. We are primarily interested in the LRU failure rate model which will predict the total failures over some period of time under a variety of conditions. Therefore, an appropriate test would be to predict the number of LRU failures for a six month period for several different situations. The aggregate prediction can turn out much better than the micro prediction for each combination of variables.

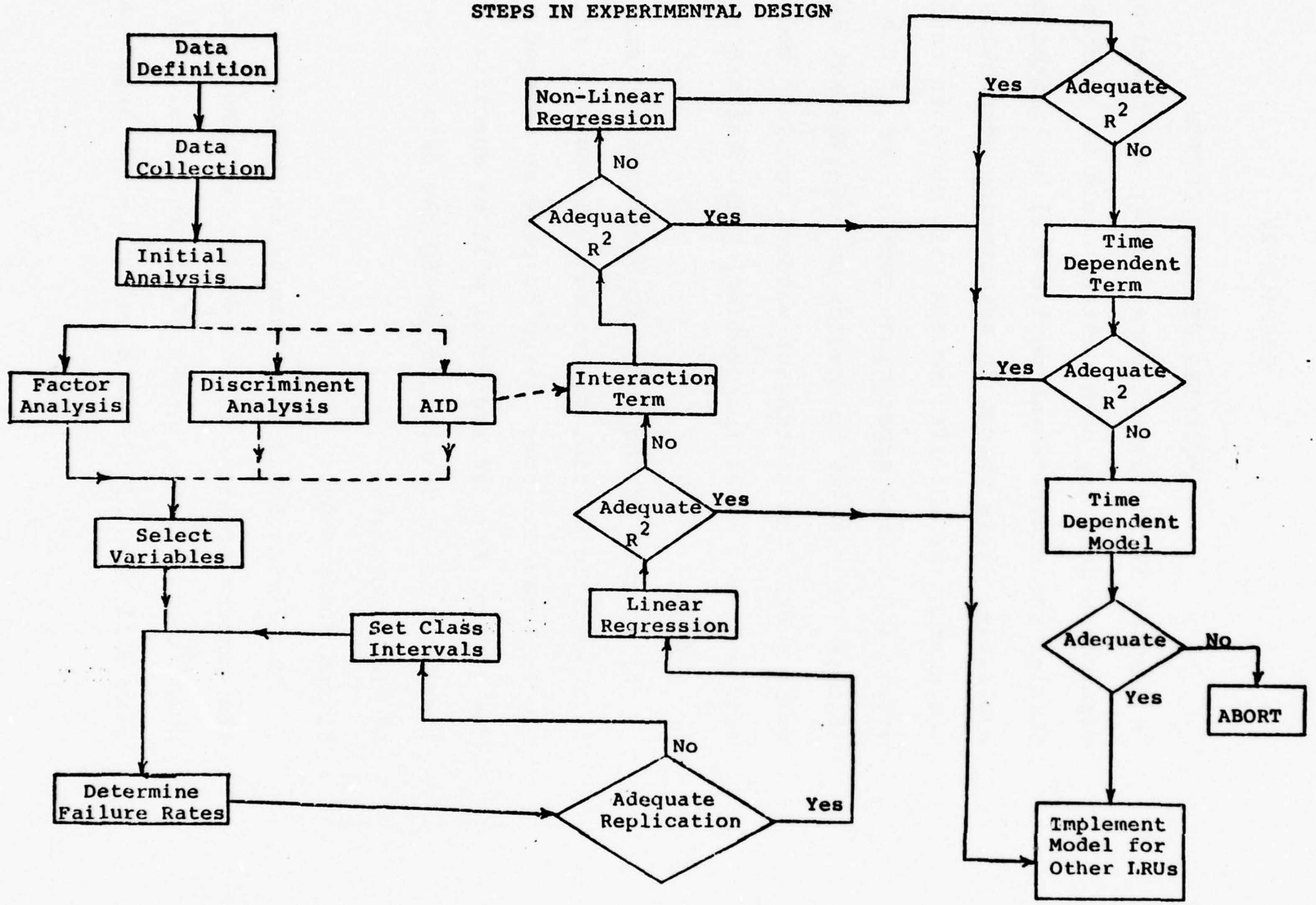
In addition to the linear regression model, a

time-dependent model should be analyzed. This could be done for sortie length versus the LRU failure rates either holding the other variables constant or by ignoring them. A model could be fit to the curve by investigating the data and fitting appropriate forms or by using the model proposed by Shurman [22]. Shurman's model would be more appropriate for a curve developed on the time to failure within a sortie, rather than the average failure rate for different length sorties however.

Figure 3.2 shows a flow diagram of the sequence of steps for analyzing the data. In Chapter IV a design for testing specific models based on their predictions of the number of LRU failures will be presented.

FIGURE 3.2

STEPS IN EXPERIMENTAL DESIGN



CHAPTER IV

MODIFIED RESEARCH DESIGN

In Chapter IV the experimental design was based on the assumption that any data desired could be obtained. In this chapter the experimental design will be limited to the data available. This change in orientation generally results in the testing of specific models with the data rather than the derivation of the appropriate model. One of the objectives of this project was to develop an experimental design which would permit the testing of several specific models to determine a "good" model for predicting LRU failure rates.

1. Models

Since most of the data are aggregate and cannot be traced to a specific sortie, the models will have to relate to total LRU failures for some period of time as opposed to a sortie. The general form of the model will be specified with the constants to be determined during the data collection phase of the experiment.

Flying Hours Model:

The flying hours model can be formulated in several ways. The current practice is to estimate the LRU failures as directly proportional to the number of flying hours. The model specified here will be a little more general -- a linear model:

(1) $F_t = a+b \cdot H_t$

where: F_t is the total LRU failures during period t.

a is a constant term representing the number of LRU failures to maintain any level of activity.

b is the constant LRU failure rate per flying hour.

H_t is the total flying hours during period t.

Sortie Model:

In the sortie model the LRU failure rate is a function of the number of sorties flown. This assumes a strong relationship between the LRU failure rate and the number of take-offs and landings. Again a linear model will be specified.

(2) $F_t = a+b \cdot S_t$

where: F_t and a are as above.

b is the number of LRU failures per sortie.

S_t is the number of sorties flow during period t.

Flying Hours and Sortie Model:

Since flying hours and sorties are both related to the LRU failures, this model will form a linear combination of both of these variables. A similar model in Casey [7] was discussed in Chapter II. The model was reported as having a R^2 of 0.667 for all maintenance write-ups for the C-5 aircraft. However, Casey did not compare the performance of the model to other models, and he only identified actions done to two digit WUCs. He used maintenance actions as the dependent variable

rather than LRU failures and he did not report on the statistical significance of the relationship. The objective here is to compare the current model to a number of commonly proposed models and test for the model which performs better than the others.

(3) $F_t = a+b\cdot H_t + c\cdot S_t$

where: F_t, a, Hz, and S_t are as above.

b is the LRU failure rate associated with flying hours.

c is the LRU failure rate associated with a sortie.

This model could be particularly appealing since the sortie coefficient could account for the high initial failures and the landing induced failures while the flying hours coefficient could account for the steady-state in-flight LRU failure rate.

Time Dependent Failure Rate Model:

A number of reports have contained data that indicate a strong relationship between the failure rate and the time into a sortie. There are at least two difficulties with this model. First, the relationship has been based on failures or unscheduled maintenance and not on LRU failures, and secondly, the model makes an assumption as to how the failures occur and not with when the failures are discovered. The model requires assuming that the vast majority of the failures are flight related and are not a function of time on the ground. There

has been some evidence indicating an increasing number of failures when the time between sorties increases Ryon, et. al. [31]. The form and the variables affecting the relationship that exists for LRU failures has not yet been determined. There are a large number of time dependent models that could be proposed; however, since Shurman's model [32, p. 13] has done well for some empirical comparisons. We will restrict our attention to it. Shurman's model for the probability of a failure at time t is given by equation (4):

$$(4) \qquad \lambda(t) = \frac{0.45\,\lambda_o\,T}{t + 0.08}$$

where: $\lambda(t)$ is the probability of a failure at time t.

λ_o is the in-flight steady-state failure rate.

T is the nominal unrefueled sortie length for which the aircraft was designed (in hours).

t is the time into a sortie (in hours).

For all of the above models there are constants (coefficients) which must have numerical values before the models can be evaluated and compared. These constants can be specified by individuals with a knowledge of the LRU failure rate process. However, since this process is apparently complex and it has not been studied in this particular context, guessing the constants would probably not prove fruitful. Another procedure would be to take a few samples and use them to set the values of the constants or to determine a range of values to be

evaluated (more than one equation for each type of model). The number of samples required for each model is a function of the number of constants to be estimated. Since a particular sample can be used in setting the constants for all of the models, the minimum number of samples required would be three for the flying hours and sortie model. Least squares regression can be used for determining the constants if the number of samples exceeds the number of constants to be specified.

The determination of λ_o for the time dependent model requires some additional calculations to be made. If more than one sample is used, the value of λ_o can be calculated for each sample and then averaged or the samples can be aggregated and the λ_o determined. To do this requires that the average LRU failure rate per flying hour be determined.

$$(5) \quad \bar{\lambda} = \int_o^{T_1} \frac{0.45\,\lambda_o T}{t + 0.08}\,dt \Big/ \int_o^{T_1} dt$$

$$= \frac{0.45\,\lambda_o T\ \ln\,(1+T_1/0.08)}{T_1}$$

where: $\bar{\lambda}$ is the average number of failures per flying hour (given a relatively constant sortie length).

T_1 is the average sortie length.

To apply equation (5) the average LRU failures per flying hour ($\bar{\lambda}$) must be determined from a sample of data by dividing the total number of LRU failures by the total flying hours for

the sample. The average sortie length (T_1) can be found by dividing the total flying hours by the total number of sorties. T is the nominal unrefueled sortie length for which the aircraft was designed and it can be estimated from the aircraft specification, imputed from a set of composite data (available in [5]), or set equal to T_1 if it is felt that T_1 is not radically different from T. Therefore, λ_o is the only unknown quantity in (5) and can be solved for as shown in (6).

$$(6)\quad \lambda_o = \frac{T_1 \cdot \bar{\lambda}}{0.45\ T\ \ln\ (1+T_1/0.08)}$$

Now by substituting for T_1 the various sortie lengths into equation (7), the expected number of LRU failures per sortie can be calculated. By multiplying this value times the number of sorties for each sortie length, an estimate of the total LRU failures can be obtained. If the particular mix at various sortie lengths is unknown, the calculation can be performed for the average sortie length and then multiplied times the total number of sorties.

$$(7)\quad E(F_{T_1}) = 0.45\ \lambda_o T\ \ln\ (1+T_1/0.08)$$

Additional comments concerning the data will be made in the following section.

2. Data Collection.

The same parameters guide the selection of the LRU to be studied in this case as were suggested in the ideal design.

The LRU should be unique to an aircraft to permit aggregate data sources to be utilized in the analysis. It should be of relatively high cost and have a high LRU failure rate to be of interest. The high LRU failure rate is also required because of sample size considerations.

Another requirement is based on the objective of providing a method of predicting the LRU failure rate for radically changing conditions. The ideal situation would be where a particular aircraft is located at four or more bases essentially flying the same type of mission but having different sortie lengths and different total flying hours. Note: Care should be taken in handling the data for an LRU when the LRU selected appears on the aircraft in multiples, i.e., multiple engine aircraft.

Part of the data required for this analysis is available on base data tapes, ABD64A. These records are six month cumulative data tapes which contain the aircraft mission-design-series (MDS) code with the data. The TRN9T07 program extracts the data for the desired MDS code and writes it on seven track tapes for further processing. The GETDATA program takes the data on the seven track tapes and records it by WUC and type of maintenance action putting the code R on all remove and replace actions for repairable items. The COLLECT program accumulates and sums manhours, total elapsed time, number of removals, and number of maintenance actions by WUC.

The number of sorties and flying hours flown covering the same period as the ABD64A tape can be obtained from base level K-18 reports or G033B-NW1A summary of sorties and flying hours by base from the AFM 65-110 data system. These programs and the steps necessary to obtain the data are described in AFHRL-TR-74-97(III)[34].

One advantage of isolating the information by base is that the six month period generally covers several thousand sorties and flying hours. This gives a sufficient sample size to be reasonably confident of the number of LRU failures for a relatively high LRU failure rate. A second advantage is that a sample of data can be obtained for each base with aircraft flying a similar mission during the same six month period. Therefore, two years of data for six bases would result in twenty-four samples.

3. Data Analysis

Once the data are collected and the estimates of the LRU failures made by each of the four models, the results can be compared. Since each model has to be compared to the actual results for each sample, the data then becomes a series of differences between the actual LRU failures and the forecasted LRU failures for each model, i.e., errors. The average error for each model will be affected by the sign of the individual error terms and positive errors will offset negative errors. Thus, a model with a low average error may predict the

aggregate LRU failures for all of the samples but not do that well on any particular sample. Therefore, the absolute errors will be used to determine an average absolute error, the mean absolute deviation (MAD). The MAD gives a measure of the average amount the forecast missed the actual value, but it does not indicate rather the forecast was over or under the actual value.

There are several ways to analyze these results. The lowest MAD for any of the models could be determined and that model selected. The more appropriate procedure would be to statistically test if any one model was significantly better than any other model. There are several ways to test this. The first two ways we will consider require the assumption that the errors are normally distributed. Then an assumption concerning equal or not equal variances must be made. If the populations are normal there is an approximate relationship between the MAD and the standard deviation. Therefore, if the MADs are different, the assumption of equal variances for this case would not appear to be valid. However, it would appear that for equal MADs with unequal variances, the model with the more predictable MAD, i.e., the one with the smaller variance, would be preferable.

The hypotheses will have to be stated for pairs of models with the model with the higher MAD being rejected in favor of the model with the lower MAD.

H_o: $MAD_1 = MAD_2$

H_a: $MAD_1 = MAD_2$

where: H_o stands for the null hypothesis.

H_a stands for the alternate hypothesis.

In testing the null hypothesis the actual test is performed on the differences between the two mean values to see if it is significantly greater than zero. The test procedure for equal sample sizes when $\sigma_1^2 = \sigma_2^2$ is:

Calculate: $t = (MAD_1 - MAD_2)/[(S_1^2 + S_2^2)/n]^{1/2}$

Accept: H_o for $-t_{1-\alpha/2,2n-2} < t < t_{1-\alpha/2,2n-2}$

Reject H_o otherwise

where: S_1^2 is the variance of the absolute deviations about the MAD_1 for the 1th model.

t is the calculated value for the student's t distribution.

α is the specified probability of rejecting H_o when H_o is true.

The test procedure for equal sample sizes when $\sigma_1^2 \neq \sigma_2^2$ is:

Calculate: $t = (MAD_1 - MAD_2)/[(S_1^2 + S_2^2)/n)^{1/2}$

Accept: H_o for $-t_{1-\alpha/2,\ n-1} < t < t_{1-\alpha/2,\ n-1}$

Reject: H_o otherwise

The tests given above assume no knowledge of which models will give better results. If the results of the samples are considered, the lowest MAD could be tested against the others using a one-tailed test. For additional information concerning these tests see Ostle [29, pp. 119-122].

If we are unwilling to assume that the differences are normally distributed, a distribution-free test can be used. Either the sign test or the signed rank test (sometimes referred to as the Wilcoxson signed rank test) can be used. Since the latter test is more efficient and the differences between the models do have magnitudes, the signed rank test should be applied. In this case the absolute errors for each sample must be paired for two of the models used to predict LRU failures. Then subtract the set of differences for one model from the differences for the other model preserving the sign of the difference (plus or minus). These values are then assigned a rank starting with the smallest and going to the highest. In ranking these values, the signs of the differences are ignored and tied values are assigned average ranks. Next the ranks are put into two columns based on whether the value of the difference has a positive or negative sign. The two columns of ranks are summed and T is set equal to the smaller of the absolute values of the two sums. If the observed T is less than or equal to the tabulated value

for the appropriate significance level, the null hypothesis of no difference between the two models is rejected. For a specific example of how to apply the signed rank test see Ostle [29, pp. 467-469].

Once the test have been run, one of these models may prove to be better than the others. If not, the information gained from doing the analysis may provide insight into developing additional models for testing. If more samples are accumulated, the tests can be continued to reduce the error of saying the null hypothesis is true when it is not (rejecting the alternate hypothesis when it is true). In addition to comparing the models, an estimate of the size of the error incurred by employing one of these models has been obtained. The square root of the variance term for the MAD for each model can be used as an estimate of the standard deviation of the forecast error. With the assumption of the errors being normally distributed, confidence limits can be established for the estimates of LRU failures. This result may prove valuable enough to make the experiment worthwhile.

CHAPTER V
SUMMARY AND RECOMMENDATION

1. Summary

In Chapter III the ideal experimental design was based on the assumption that there would be no restriction on the type or quantity of data available. For the ideal design most of the variables given as explanations of variations in failure types of data were listed in Chapter II. In addition, some other variables thought to affect the failure process were listed as well as some variables mentioned that are felt to affect the failure process but which do not appear readily quantifiable. The objective of the ideal design is to determine the variables with a significant impact on the LRU failure rate and the form of the model which translates these variables into the number of LRU failures when the basic policies of flying hours, sorties and mission types change.

There is a number of problems associated with this design. The data are not available and the system for collecting such data would need to be developed. The quantity of data required is large requiring considerable time, effort, and cost. The analysis of the data (given a complete ANOVA is not possible) is not as all-inclusive as would be liked. Finally, there is no guarantee that a model that will do a good or even an adequate job of predicting LRU failures will result from the analysis.

There is a number of advantages associated with performing the analysis however. The strength and form of the relationship

between a number of variables and the LRU failure rate could be determined. This would point the way to what types of data should be collected and which types of data should not be collected. The type of the relationship existing between the variables and the LRU failure rate could affect maintenance policies as well as inventory policies. The error in using these models could be estimated from the analysis providing confidence limits for the forecasts. The data collected could also be used for a variety of other studies which have not been possible to do because currently the current data are not tractable to the specific sortie and aircraft.

In Chapter IV the experimental design was based more on the concept of what can be done with the available data. This requires using an aggregate model since the data available are primarily aggregate data. A causal model is less appealing and the problem becomes more of a curve fitting problem, i.e., fit the data with an equation which does a good job of predicting but may not explain the underlying relationships.

There are several problems associated with this approach. The one already noted of not being particularly helpful in understanding the underlying process or relationships. Since the model is not a causal one, predicting for conditions outside the parameters on which the model as tested becomes very risky. In addition, by testing only a few specific models, the "best of the lot" may be poor model and a good model may be missed because of the poor specification of the initial constants.

The advantages all relate to the minimal requirements for data, time, and effort. Several models can be compared in a relatively short time and the "best of the lot" picked. If it is better than what you are currently doing, then an improvement has been made. The data currently are available which shortens the time frame for doing the analysis and the cost of obtaining the data. In addition, the time dependent failure rate model has considerable creditability and should receive some immediate attention. A study of Shurman's model on some empirical data could speed the process of implementation of time dependent models if appropriate, or by the same token, the study could indicate that a more detailed, long-term analysis of time dependent models should be undertaken.

2. Recommendation

The recommendation concerning the prediction of LRU failure rates is to implement the design outlined in Chapter IV. The design is such that other models using the same variables could easily be added to the analysis if it becomes desirable. If a model can be determined which does a satifactory job in predicting the LRU failure rate, the impetus for doing the analysis described in Chapter III would diminish. However, the recommendation would be to work towards the recording of the type of data described in Chapter III. There is a tremendous effort already being expended on data collection. The major problem with the current system is the inability to trace the data back once the data has been aggregated. The additional

effort required to record the data appropriately may be minimal and the change relatively easy. Then the potential for a number of studies affecting policies in inventory, maintenance, etc. could be done. The end result might benefit the data collection system by determining which data are useful for decision making purposes and which data should not be collected at all.

REFERENCES

1. ARINC Research Corporation, "Development of Flight Safety Prediction Methodology for U. S. Naval Safety Center," Final Report, Pub. No. 753-01-3-982 (Rev. 1), ARINC Research Corporation, Western Division, Santa Ana, California, February 1970.

2. ARINC Research Corporation, "Flight Safety Prediction Techniques Feasibility Study," Pub. No. 705-01-1-777, ARINC Research Corporation, Western Division, Santa Ana, California, May 1967.

3. ARIND, "Final Report, F-106 UHF Flight Test Evaluation," ARIND Report No. 329-01-3-499, June 1965, pp 11 and 19.

4. Bell, C. F. and T. C. Smith, "The Oxnard Base Maintenance Management Improvement Program," The RAND Corporation, RM-3370-PR, November 1962.

5. Boeing Company, "Aircraft Field Experience, Composite Handbook," Document D6-53583, 1969 & 1974.

6. Boeing Company, "B-52D Operations - Southeast Asia vs CONUS," Document D 162-10015-1, Sep 1970.

7. Casey, N. L., "Logistics Impact of Longer C-5 Missions," AFLC Technical Report No. 76-16, May 1977.

8. Cooley, W. W. and P. R. Lohnes, "Multivariate Procedures for the Behaviorial Sciences," New York, Wiley, 1962.

9. Dantowitz, Allan, George Hirschberger, and David Pravidlo, "Analysis of Aeronautical Equipment Environmental Failures," Technical Report #AFFDL-TR-71-32, May 1971.

10. Donaldson, T. S. and Major G. W. Burke, "Aircraft Performance Related to Ground Alert Duration," The RAND Corporation, RM-4628-PR, March 1966 (SECRET).

11. Donaldson, T. S. and A. F. Sweetland, "Trends in F-101 Aircraft Maintenance Requirements," The RAND Corporation, RM-4930-PR, April 1966.

12. Donaldson, T. S. and A. F. Sweetland, "The Relationship of Flight-Line Maintenance Man-Hours to Aircraft Flying Hours," The RAND Corporation, RM-5701-PR, August 1968.

13. Dow, Charles L. and Walter L. Schnee, "A Study of Flying Hours and Sortie as Predictors of B-52H Engine Failures," AFIT Masters Thesis, January 1972.

14. Green, Paul E. and Donald S. Tull, Research for Marketing Decisions, Englewood Cliffs, N. J.; Prentice-Hall, 1966.

15. Harmon, Harry H., Modern Factor Analysis, 2nd Edition, University of Chicago Press, 1967.

16. Horn, R. L., and G. S. Shoup, "Determination and Use of Failure Patterns," Proc. 8th National Symposium on Reliability and Quality Control, Washington, D. C., January 9-11, 1962, pp 416-424.

17. Hunsaker, H. D., F. Conway, and E. R. Doherty, "Logistics Impact of Sortie Oriented Crew Training," AFLC Technical Report No. 76-7, April 1977.

18. Johnson, W. L., and K. E. Reed, "F-4D Field Experience Part IV Summary," AFFDL-TR-72-2, 17 January 1972.

19. La Vallee, R. S. and D. S. Stoller, "Factors Affecting Malfunction Rates for F-86F and F-86D Aircraft," The RAND Corporation, RM-1790, September 12, 1956.

20. McCabe, E. L., W. D. Pearce and G. D. Rice, "Method for Developing Equipment Failure Rate K Factors," D 180-17674-2, Boeing Company, January 3, 1975.

21. McGlothlin, W. H. and T. S. Donaldson, "Trends in Aircraft Maintenance Requirements," The RAND Corporation, RM-4049-PR, June 1964.

22. Martinez, Edward, "Study on 'When Discovered' Codes for Math Modeling," Service Engineering Division, Directorate of Materiel Management, San Antonio AMA, Kelly AFB, Texas, September 1968.

23. Martz, H. F. Jr., K. Campbell, and H. T. Davis, "Estimating and Forecasting Failure-rate Processes by Means of the Kulman Filter," Theory and Applications of Reliability, Vol 1, Academic Press, 1977.

24. Morgan, J. N. and J. A. Sonquist, "Problems in the Analysis of Survey Data and a Proposal," Journal of the American Statistical Association, 58 (June 1963), 415-35.

25. Morgan, J. N. and J. A. Sonquist, "The Detection of Interaction Effects," The University of Michigan Institute for Social Research, Monograph No. 35, 1964.

26. Morrison, Donald F., Multivariate Statistical Methods, New York, McGraw-Hill Book Company, Inc., 1967.

27. Morrison, Donald G., "On the Interpretation of Discriminant Analysis," Journal of Marketing Research, Vol. VI, May 1969, pp 156-163.

28. Neter, John and William Wasserman, Applied Linear Statistical Models, Richard D. Irwin, Inc., Homewood, Illinois, 1974.

29. Ostle, Bernard, Statistics in Research: Second Edition, Iowa State University Press, Ames, Iowa, 1964.

30. Rummel, R. J., "Understanding Factor Analysis," Journal of Conflict Resolution, Vol. XI No. 4, pp 444-80.

31. Ryon, Commander G. G., Major J. B. Abell and Major F. D. Livers, "Air Analysis of Factors Influencing Intercepter Aircraft Break-Rates," AFIT Thesis, August 1968.

32. Shurman, Maurice B., "Time-Dependent Failure Rates for Jet Aircraft," report released January 30, 1978, Boeing Co.

33. Steiner, R. G., "Spares Consumption Rate vs. Sortie Length for Airframes, Avionics, and Engines of Boeing Commercial Jet Aircraft," The Boeing Co., December 1966.

34. Tetmeyer, Major Donald E., Sharon R. Nichols and Robert N. Deem, "Simulating Maintenance Manning for New Weapons Systems: Maintenance Data Analysis Programs," AFHRL-TR-74-97 (III), May 1976.

35. Tomlin, Eric M., "Failure Prediction Techniques Applied to the C-47 Aircraft Engine," Report #USAMC-LTC-3-73-18, May 1973.